

# Quasi-sure convergence theorem in $p$ -variation distance for stochastic differential equations

H. Boedihardjo\* and Z. Qian†

March 3, 2013

## Abstract

In this paper by calculating carefully the capacities (defined by high order Sobolev norms on the Wiener space) for some functions of Brownian motion, we show that the dyadic approximations of the sample paths of the Brownian motion converge in the  $p$ -variation distance to the Brownian motion except for a slim set (i.e. except for a zero subset with respect to the capacity on the Wiener space of any order). This presents a way for studying quasi-sure properties of Wiener functionals by means of the rough path analysis.

## 1 Introduction

It has been suggested that the recent theory of rough paths, put forward in T. Lyons [36], and developed further over the past years by T. Lyons with his coauthors, and other authors, see [35], should simplify and strengthen the results in quasi-sure analysis over the Wiener space, which was initiated by P. Malliavin (see [40] for example). In fact, the rough path analysis has direct applications in solving stochastic differential equations quasi-surely (see below for a precise meaning). What however is missing in literature is an approximation theorem towards Brownian motion sample paths by simple random curves in  $p$ -variation distance and in quasi-sure sense, i.e. except for a slim subset, instead of a probability zero set, see below for a definition of slim subsets in Wiener space. The main goal of the paper is to establish such a theorem (see Theorem 2.3 and Theorem 2.4 below). The results obtained in the paper allow to study quasi-sure properties for important Wiener functionals – solutions of Itô's stochastic differential equations. The main step in our proof is the construction of quasi-surely defined *geometric* rough paths associated with Brownian motion, which we believe has other applications, though not discussed in the present paper. It is known that there is an equivalence of capacity zero sets (defined in terms of Dirichlet norms of the Ornstein-Uhlenbeck process on the Wiener space) and the polar sets defined by the Brownian sheets. Therefore, with the quasi-surely defined geometric Brownian motion paths we constructed,

---

\*Mathematical Institute, University of Oxford, Oxford OX1 3LB, England.

Email: Horatio.Boedihardjo@maths.ox.ac.uk

†Mathematical Institute, University of Oxford, Oxford OX1 3LB, England. Email: qianz@maths.ox.ac.uk.

it is possible to define the rough path analysis for the Brownian sheet, thus provides another possible route to study stochastic partial differential equations via the rough path analysis, this potential application however is not pursued further here. There is increasing interest in applying rough path analysis to the study of stochastic partial differential equations, such as the recent papers [2], [4], [5], [9], [19], [20] and the references therein.

In order to address the question we investigate in this perspective, we begin with some elements in the analysis of rough paths, and establish the notions and notations which will be used throughout the paper.

## 1.1 Concept of rough paths

Let  $(\mathbb{R}^d)^{\otimes k} = \mathbb{R}^d \otimes \cdots \otimes \mathbb{R}^d$  be the tensor product of  $k$ -folds of the Euclidean space  $\mathbb{R}^d$ ; which may be identified with  $\mathbb{R}^{kd}$ ; equipped with the corresponding Euclidean norm. For a positive integer  $n$ ,  $T^{(n)}(\mathbb{R}^d)$  denotes the truncated tensor algebra

$$T^{(n)}(\mathbb{R}^d) = \sum_{k=0}^n \oplus (\mathbb{R}^d)^{\otimes k},$$

where  $(\mathbb{R}^d)^{\otimes 0} = \mathbb{R}^1$ .

Given  $T > 0$ ,  $\Delta = \{(s, t) : 0 \leq s \leq t \leq T\}$ .  $T > 0$  will be fixed but arbitrary, so it will be assumed to be 1 without lose of generality.

A *continuous* path  $w : [0, T] \rightarrow \mathbb{R}^d$  is said to have finite total variation on  $[0, T]$ , if

$$\sup_D \sum_l |w_{t_l} - w_{t_{l-1}}| < \infty$$

where  $\sup_D$  takes over all finite partitions of  $[0, T]$ :

$$D = \{0 = t_0 < \cdots < t_m = T\}.$$

This convention applies to similar situations below without further qualification.

Let  $\Omega^\infty(\mathbb{R}^d)$  denote the totality of all continuous paths in  $\mathbb{R}^d$  with finite total variations on  $[0, T]$ . If  $w \in \Omega^\infty(\mathbb{R}^d)$ , the  $k$ -th iterated path integral over  $[s, t]$

$$w_{s,t}^k = \int_{s < t_1 < \cdots < t_k < t} dw_{t_1} \otimes \cdots \otimes dw_{t_k}.$$

By definition,  $w_{s,t}^1 = w_t - w_s$  is the increment of the path  $w$  over  $[s, t]$ , and for  $k \geq 2$

$$w_{s,t}^k = \lim_{m(D) \rightarrow 0} \sum_l \sum_{j=1}^{k-1} w_{s,t_{l-1}}^j \otimes w_{t_{l-1},t_l}^{k-j}$$

are defined inductively. Collecting all  $k$ -th iterated integrals (up to degree  $n$ ) together we define  $L_n(w) : \Delta \rightarrow T^{(n)}(\mathbb{R}^d)$  by

$$L_n(w)_{s,t} = (1, L_n(w)_{s,t}^1, \cdots, L_n(w)_{s,t}^n), \quad L_n(w)_{s,t}^k = w_{s,t}^k \quad \forall (s, t) \in \Delta$$

$L_n(w)^1$  is called the first level path of  $L_n(w)$  which indeed recovers the original path through  $w_t = w_0 + L_n(w)_{0,t}^1$  up to the starting point.  $L_n(w)^2$  is called the second level path etc. Often  $L_n(w)$  is written as  $\mathbf{w}$  if no confusion is possible.

$L_n(w)$  satisfies an important equation, called Chen's identity

$$L_n(w)_{s,r} \otimes L_n(w)_{r,t} = L_n(w)_{s,t} \quad \forall 0 \leq s < r < t,$$

where the tensor product takes place in the truncated tensor algebra  $T^{(n)}(\mathbb{R}^d)$ . It is indeed the reason that the zeroth term is taken as 1 in the definition of  $L_n(w)$ . Chen's identity is nothing but represents the additivity of iterated integrals over disjoint intervals.

Let  $\Omega^{\infty,n}(\mathbb{R}^d)$  denote the totality of all functions  $L_n(w)$  where  $w$  runs through  $\Omega^\infty(\mathbb{R}^d)$ :

$$\Omega^{\infty,n}(\mathbb{R}^d) = L_n(\Omega^\infty(\mathbb{R}^d)) = \{L_n(w) : w \in \Omega^\infty(\mathbb{R}^d)\}$$

which may be naturally identified with the space of all  $w \in \Omega^\infty(\mathbb{R}^d)$  started from 0 (or any fixed point in  $\mathbb{R}^d$ ).

Next step is to equip  $\Omega^{\infty,n}(\mathbb{R}^d)$  with a metric, and introduce the concept of geometric rough paths in  $\mathbb{R}^d$ . Let  $p \geq 1$  be fixed and  $[p]$  denote the integer part of  $p$ , which relates to the roughness of *sample paths*. The interesting values of  $p$  are real numbers between 2 and 3 for the study of Brownian motion in  $\mathbb{R}^d$ . The  $p$ -variation metric, which is the key concept in the analysis of rough paths, denoted by  $d_p$ , is a metric on  $\Omega^{\infty,[p]}(\mathbb{R}^d)$  defined by

$$d_p(L(v), L(w)) = \max_{1 \leq k \leq [p]} \sup_D \left( \sum_l \left| L(v)_{t_{l-1}, t_l}^k - L(w)_{t_{l-1}, t_l}^k \right|^{\frac{p}{k}} \right)^{\frac{k}{p}}$$

where  $L(w)$  denotes  $L_{[p]}(w)$  for simplicity. The completion of  $\Omega^{\infty,[p]}(\mathbb{R}^d)$  under  $d_p$  is denoted by  $G\Omega_p(\mathbb{R}^d)$ . An element in  $G\Omega_p(\mathbb{R}^d)$  is called a *geometric rough path* in  $\mathbb{R}^d$  of roughness  $p$ .

If  $\mathbf{w} = (1, w^1, \dots, w^{[p]}) \in G\Omega_p(\mathbb{R}^d)$ , then  $\mathbf{w}$  satisfies Chen's identity  $\mathbf{w}_{s,r} \otimes \mathbf{w}_{r,t} = \mathbf{w}_{s,t}$  in  $T^{[p]}(\mathbb{R}^d)$  (for any  $0 \leq s < r < t \leq T$ ), and  $\mathbf{w}$  has finite  $p$ -variation in the sense that  $\sup_D \sum_l \left| w_{t_{l-1}, t_l}^k \right|^{p/k} < \infty$  for all  $k \leq [p]$ .

T. Lyons [36] has demonstrated that a theory of integration for a geometric rough path may be established. Let  $\mathbf{w} \in G\Omega_p(\mathbb{R}^d)$  and  $f : \mathbb{R}^d \rightarrow L(\mathbb{R}^d, \mathbb{R}^{\tilde{d}})$  a function on  $\mathbb{R}^d$  with values in the linear space  $L(\mathbb{R}^d, \mathbb{R}^{\tilde{d}})$  of all linear operators from  $\mathbb{R}^d$  to  $\mathbb{R}^{\tilde{d}}$ , where  $L(\mathbb{R}^d, \mathbb{R}^{\tilde{d}})$  may be identified with  $\mathbb{R}^d \otimes \mathbb{R}^{\tilde{d}}$  or the Euclidean space  $\mathbb{R}^{d\tilde{d}}$ . Such an  $f$  is called an  $\mathbb{R}^{\tilde{d}}$ -valued 1-form on  $\mathbb{R}^d$ . Let  $f^k$  denote the  $k-1$ -th derivative  $D^{k-1}f$  of  $f$  which is identified with a function on  $\mathbb{R}^d$  valued in  $L((\mathbb{R}^d)^{\otimes k}, \mathbb{R}^{\tilde{d}})$  (where  $k = 1, \dots$ ). In particular  $f^1 = f$ .

If  $w \in \Omega^\infty(\mathbb{R}^d)$ , then  $y = (y_t)_{t \in [0, T]} \in \Omega^\infty(\mathbb{R}^{\tilde{d}})$  where  $y_t = \int_0^t f(w_s) dw_s$  is the path integral defined via the Riemann integral

$$y_t = \lim_{m(D) \downarrow 0} \sum_l f(w_{t_{l-1}})(w_{t_l} - w_{t_{l-1}}). \quad (1.1)$$

One of the main results in the rough path analysis is the following *continuity theorem*.

**Theorem 1.1** (*T. Lyons [36]*) Suppose that  $f \in C_b^{n+1}(\mathbb{R}^d; L(\mathbb{R}^d, \mathbb{R}^{\tilde{d}}))$ , where  $n \geq 1$  is an integer, and suppose that  $p \geq 1$  such that  $[p] \leq n$ . Then the integration, which takes  $L(w) \in \Omega^{\infty, [p]}(\mathbb{R}^d)$  to the lifting  $L(y) \in \Omega^{\infty, [p]}(\mathbb{R}^{\tilde{d}})$  of  $y$  defined by (1.1), is continuous with respect to the  $p$ -variation metrics. Moreover, the mapping  $L(w) \rightarrow L(y)$  is uniformly continuous on any bounded set of  $\Omega^{\infty, [p]}(\mathbb{R}^d)$ .

If  $w \in \Omega^\infty(\mathbb{R}^d)$  and its lifting to a geometric rough path  $\mathbf{w} = (1, w^1, \dots, w^{[p]})$ ,  $y_t = \int_0^t f(w_s)dw_s$  the usual Riemannian integral defined as above, then  $y \in \Omega^\infty(\mathbb{R}^{\tilde{d}})$  (which is true actually for a Lipschitz continuous  $f$ ). The lifting  $\mathbf{y} = L_{[p]}(y)$  is denoted by  $\int f(\mathbf{w})d\mathbf{w}$ . The previous theorem says that the *Itô-Lyons integration*  $\mathbf{w} \rightarrow \int f(\mathbf{w})d\mathbf{w}$  is continuous with respect to the  $p$ -variation metrics. Notice that the usual path integral  $w \rightarrow \int_0^t f(w_s)dw_s$  is in general not continuous under the uniform norm of paths.

As a consequence, for  $\mathbf{w} \in G\Omega_p(\mathbb{R}^d)$  and  $f \in C_b^{[p]+1}(\mathbb{R}^d; L(\mathbb{R}^d, \mathbb{R}^{\tilde{d}}))$  we can definite its integral  $\int f(\mathbf{w})d\mathbf{w}$  as a unique geometric rough path in  $G\Omega_p(\mathbb{R}^{\tilde{d}})$ . It is however interesting to know how to define  $\int f(\mathbf{w})d\mathbf{w}$  directly by means of rough paths.

Let us describe the definition for  $p \in (2, 3)$  which is the most interesting case as it is the case for geometric rough paths associated with Brownian motion.

Since  $[p] = 2$ , so that we need to define two components  $y^1, y^2$  which defines a rough path  $\int f(\mathbf{w})d\mathbf{w} \equiv (1, y^1, y^2)$  where  $\mathbf{w} = (1, w^1, w^2) \in G\Omega_p(\mathbb{R}^d)$  and  $f \in C_b^3(\mathbb{R}^d; L(\mathbb{R}^d, \mathbb{R}^{\tilde{d}}))$ . To this end first define  $\tilde{\mathbf{y}} = (1, \tilde{y}^1, \tilde{y}^2)$  by

$$\tilde{y}_{s,t}^1 = f^1(w_s^1)(w_{s,t}^1) + f^2(w_s^1)(w_{s,t}^2)$$

and

$$\tilde{y}_{s,t}^2 = f^1(w_s^1) \otimes f^1(w_s^1)(w_{s,t}^2).$$

It is understandable that  $\tilde{\mathbf{y}}$  is not a geometric rough path yet (for example, it does not satisfy Chen's identity in general), so we take a limiting procedure to define  $\mathbf{y} = (1, y^1, y^2)$  by means of Riemann sums but at the tensor level. More precisely define

$$\mathbf{y}_{s,t} = \lim_{m(D) \downarrow 0} \mathbf{y}_{t_0, t_1} \otimes \dots \otimes \mathbf{y}_{t_{m-1}, t_m}$$

where the tensor product  $\otimes$  takes place in the truncated algebra  $T^2(\mathbb{R}^{\tilde{d}})$ , and the limit  $\lim_{m(D) \downarrow 0}$  takes over finite partitions of  $[s, t]$ . We then can show that the above limit exists and  $\mathbf{y} = \int f(\mathbf{w})d\mathbf{w}$ .

## 1.2 Differential equations driven by rough paths

The most important result in the rough path analysis is the universal limit theorem for solutions of differential equations. Let  $\mathbb{R}^d$  and  $\mathbb{R}^{\tilde{d}}$  be two Euclidean spaces. Consider a system of differential equations of the following form

$$dy_t^j = \sum_{i=1}^d f_i^j(y_t)dw_t^i \quad (1.2)$$

with initial data  $y_0 \in \mathbb{R}^{\tilde{d}}$ ,  $w = (w^i)$  is a continuous path in  $\mathbb{R}^d$ , and  $f = (f_i^j) : \mathbb{R}^{\tilde{d}} \rightarrow L(\mathbb{R}^d, \mathbb{R}^{\tilde{d}})$  is a function, called an  $\mathbb{R}^d$ -valued vector field on  $\mathbb{R}^{\tilde{d}}$ , where for  $y \in \mathbb{R}^{\tilde{d}}$ ,  $\xi \in \mathbb{R}^d$ ,  $(f(y)\xi)^j = \sum_{i=1}^d f_i^j(y)\xi^i$ . Suppose  $f_i^j$  are Lipschitz continuous, and  $w \in \Omega^\infty(\mathbb{R}^d)$ , then the standard Picard iteration applying to the integral equation

$$y_t^j = y_0^j + \sum_{i=1}^d \int_0^t f_i^j(y_s) dw_s^i$$

allows to determine a unique continuous path  $y$  in  $\mathbb{R}^{\tilde{d}}$  with total finite variation. This establishes a mapping sending  $w \in \Omega^\infty(\mathbb{R}^d)$  to the solution  $y \in \Omega^\infty(\mathbb{R}^{\tilde{d}})$ , denoted by  $y = F(y_0, w)$ . We then lift both paths  $w$  and  $y$  to their corresponding geometric rough paths of the same roughness  $p$ , which thus defines a mapping which maps  $L_{[p]}(w)$  to  $L_{[p]}(y)$  for each  $p \geq 1$ . We will again denote it as  $F(y_0, \cdot)$ . That is

$$F(y_0, L_n(w)) = L_n(F(y_0, w)) \quad \forall w \in \Omega^{\infty, n}(\mathbb{R}^d).$$

**Theorem 1.2** (*T. Lyons [36]*) *Let  $p \geq 1$  and  $f = (f_i^j) \in C_b^{[p]+1}(\mathbb{R}^{\tilde{d}}; L(\mathbb{R}^d, \mathbb{R}^{\tilde{d}}))$ . Then*

$$F(y_0, \cdot) : L_{[p]}(w) \rightarrow L_{[p]}(F(y_0, w))$$

*is continuous from  $\Omega^{\infty, [p]}(\mathbb{R}^d)$  to  $\Omega^{\infty, [p]}(\mathbb{R}^{\tilde{d}})$  with respect to the corresponding  $p$ -variation distances.*

This theorem ensures that there is a unique continuous extension of  $F$  on  $G\Omega_p(\mathbb{R}^d)$ , still denoted by  $F(y_0, \cdot)$ , so that  $F(y_0, \mathbf{w}) = L_{[p]}(F(y_0, \pi(\mathbf{w})))$  if  $\mathbf{w} \in G\Omega_p(\mathbb{R}^d)$ . The continuous mapping

$$\mathbf{w} \in G\Omega_p(\mathbb{R}^d) \rightarrow F(y_0, \mathbf{w}) \in G\Omega_p(\mathbb{R}^{\tilde{d}})$$

is called the *Itô-Lyons mapping* defined by the differential equation (1.2).

## 2 Main results

The analysis of rough paths, developed in T. Lyons [36], [35], can be applied to the study of stochastic differential equations with driven noises which are far more irregular than those of sample paths of semimartingales. On the other hand, Lyons' theory also sheds new insight on Itô's classical theory of stochastic differential equations, namely, stochastic differential equations driven by Brownian motion, as presented in [24] for example.

In order to apply Lyons' rough path theory to Itô's theory of stochastic differential equations, it is necessary to enhance Brownian motion sample paths into geometric rough paths. The first construction of Brownian motion as rough paths was presented in E.-M. Sipilainen's Ph.D. thesis (1993) at University of Edinburgh, under the supervision of T. Lyons. B. Hambly and T. Lyons provide further examples in [1] and [21]. They proved that symmetric diffusions with generators of elliptic differential operators of second order, and Brownian motion on the Sierpinski gasket can be enhanced into geometric rough paths of level two. In [6], geometric rough paths of level 3 associated with fractional Brownian

motions with Hurst parameter  $h > \frac{1}{4}$  were constructed by means of dyadic approximations, and in [35] further examples of geometric rough paths of level 2 or 3 associated a class of Gaussian processes and more generally a class of stochastic processes with long time memory are given. In particular, T. Lyons and Z. Qian [35] showed that if the correlation of a continuous stochastic process over disjoint time intervals satisfies a polynomial decay condition, together with further less important technical conditions, then the rough path analysis may be applied to stochastic differential equations driven by such a process. We notice that the correlation decay condition is a generalization of the martingale property which is the key in Itô's theory of stochastic calculus. We would like to recommend the reader for other constructions of rough paths to the books by T. Lyons, M. Caruana, and T. Lévy [37], P. Friz and N. Victoir [12], A. Lejay's survey [32], and other papers listed in the references in [35] and [12].

In particular, the following theorem has been proved (for a proof see for example [35]).

**Theorem 2.1** *If  $B$  is a Brownian motion in  $\mathbb{R}^d$  on a complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ ,  $\mathbf{B} = (1, B^1, B^2)$  where  $B_{s,t}^1 = B_t - B_s$  and*

$$B_{s,t}^2 = \int_{s < t_1 < t_2 < t} \circ dB_{t_1} \otimes \circ dB_{t_2} \quad (2.1)$$

*for  $0 \leq s < t$ , defined in the sense of Stratonovich's integration, then, for any  $2 < p < 3$*

$$\mathbb{P} \{ \omega : \mathbf{B}(\omega) \in G\Omega_p(\mathbb{R}^d) \} = 1.$$

The goal of the present article is to prove a stronger result that  $\mathbf{B}$  are geometric rough paths quasi-surely. To make it more precise, we need more notions and notations on Malliavin calculus, the capacity theory on the Wiener space.

Let  $\mathbf{W}$  denote the space of all continuous paths in  $\mathbb{R}^d$  started at the origin, endowed with the topology of uniform convergence over finite time intervals. If we wish to emphasize the dimension  $d$  we will render our notations with superscript  $d$ .  $\mathcal{B}(\mathbf{W})$  denotes the Borel  $\sigma$ -algebra, which can be described in terms of the coordinate process  $(B_t)_{t \geq 0}$  on  $\mathbf{W}$ , where for each  $t \geq 0$ ,  $B_t$  is the coordinate functional on  $\mathbf{W}$  at time  $t$ . That is to say, if  $w \in \mathbf{W}$ , then  $B_t(w) = w(t)$ . In what follows, in order to avoid heavy notations,  $B_t$  may be written as  $w(t)$  for  $w \in \mathbf{W}$ ,  $w_t(w)$  or  $w(t, w)$  if no confusion may arisen. Similar convention applies to  $w$  as well:  $w$  may be considered as a typical path in  $\mathbb{R}^d$  or as the canonical coordinate process on  $\mathbf{W}$ . Let  $\mathcal{F}_t^0$  to be the smallest  $\sigma$ -algebra over  $\mathbf{W}$  such that all  $B_s$  for all  $s \leq t$  are  $\mathcal{F}_t^0$ -measurable. Then  $\mathcal{F}^0 = \sigma\{B_t : t \geq 0\}$  coincides with  $\mathcal{B}(\mathbf{W})$ .

The Wiener measure  $\mathbb{P}$  is the unique probability on  $(\mathbf{W}, \mathcal{B}(\mathbf{W}))$  which makes the coordinate process  $(B_t)_{t \geq 0}$  a standard Brownian motion. Alternatively,  $\mathbb{P}$  is the Gaussian measure on  $(\mathbf{W}, \mathcal{B}(\mathbf{W}))$  with the characteristic function

$$\int_{\mathbf{W}} e^{i\langle l, \cdot \rangle} d\mathbb{P} = e^{-\frac{1}{2}\|l\|_H^2} \quad \text{for } l \in \mathbf{W}^*$$

where  $\mathbf{W}^*$  is the dual space of  $\mathbf{W}$ , identified with a linear subspace of the Cameron-Martin space

$$\mathbf{H} = \{h \in L^2(\mathbb{R}_+, \mathbb{R}^d) | h(0) = 0 \text{ and } \dot{h} \in L^2(\mathbb{R}_+, \mathbb{R}^d)\}$$

equipped with the Hilbert norm  $|h|_{\mathbf{H}} = \sqrt{\int_0^\infty |\dot{h}(t)|^2 dt}$ . It is clear that any  $h \in \mathbf{H}$  has a continuous representation and there is a natural continuous imbedding  $\mathbf{H} \hookrightarrow \mathbf{W}$ .

If  $h \in \mathbf{H}$  then  $\xi_h$  denotes the Wiener functional on  $\mathbf{W}$  defined by Itô's integration  $\xi_h = \int_0^\infty \dot{h} \cdot dw$ . Then,  $\xi_h$  has a normal distribution  $N(0, |h|_{\mathbf{H}}^2)$ . The translation  $\tau_h : w \rightarrow w + h$  is measurable, and  $\mathbb{P} \circ \tau_h$  is equivalent to the Wiener measure  $\mathbb{P}$ , which allows us to define the Malliavin derivative of  $\xi_l$  in a direction  $h$ :

$$D_h \xi_l = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \sum_{i=1}^d \int_0^\infty \dot{l}^i d(w^i + \varepsilon h^i) = \langle l, h \rangle_{\mathbf{H}},$$

Define  $D\xi_l$  to be  $\mathbf{H}$ -valued random variable on  $\mathbf{W}$  by requiring that  $\langle D\xi_l, h \rangle_{\mathbf{H}} = D_h \xi_l$  for all  $h \in \mathbf{H}$  so that  $D\xi_l = l$  for any  $l \in \mathbf{H}$ .

If  $f$  is a smooth Schwartz function on  $\mathbb{R}^n$ , and  $l_1, \dots, l_n \in \mathbf{H}$ , then  $F = f(\xi_{l_1}, \dots, \xi_{l_n})$  is called a *smooth Wiener functional* on  $\mathbf{W}$ . The collection of all such smooth functionals is denoted by  $\mathcal{S}$ . By forcing the chain rule to define the Malliavin derivative by

$$DF = Df(\xi_{l_1}, \dots, \xi_{l_n}) = \sum_{j=1}^n \frac{\partial f}{\partial x^j}(\xi_{l_1}, \dots, \xi_{l_n}) l_j.$$

By iterating the definition we may define

$$D^2 F = \sum_{j=1}^n \frac{\partial^2 f}{\partial x^i \partial x^j}(\xi_{l_1}, \dots, \xi_{l_n}) l_i \otimes l_j$$

which is an  $\mathbf{H}^{\otimes 2}$ -valued random variable. We can define  $D^s F$  inductively, see [24] for more details.

The Sobolev norms  $\|F\|_{q,N}$  (for  $N = 0, 1, 2, \dots$ ) is equivalent to  $\sum_{j \leq N} \|D^j F\|_p$ , where

$$\|D^j F\|_q = \sqrt[q]{\mathbb{E} \{ |D^j F|_{\mathbf{H}^{\otimes j}}^q \}}.$$

The completion of smooth Wiener functionals under the norm  $\|\cdot\|_{q,N}$  is denoted by  $\mathbb{D}_N^q$ . For any  $q \geq 1$  and integer  $N \geq 0$ ,  $(\mathbb{D}_N^q, \|\cdot\|_{q,N})$  is a Banach space, called a Sobolev space over the Wiener space  $\mathbf{W}$ . The  $(q, N)$ -capacity  $\text{Cap}_{q,N}$  is a function on the collection of subsets of  $\mathbf{W}$  defined as the following.

If  $A$  is an open subset of  $\mathbf{W}$ , then

$$\text{Cap}_{q,N}(A) = \inf \{ \|u\|_{q,N} : u \in \mathbb{D}_N^q \text{ s.t. } u \geq 1 \text{ } \mathbb{P}\text{-a.e. on } A \text{ and } u \geq 0 \text{ a.s. on } \mathbf{W} \}$$

and for a general subset  $A \subset \mathbf{W}$

$$\text{Cap}_{q,N}(A) = \inf \{ \text{Cap}_{q,N}(B) : B \text{ open and } B \supseteq A \}.$$

For each pair  $(q, N)$ ,  $\text{Cap}_{q,N}$  is a capacity on  $\mathbf{W}$ , in the sense that the following hold.

1.  $0 \leq \text{Cap}_{q,N}(A) \leq \text{Cap}_{q,N}(B)$  if  $A \subseteq B$ ,

2.  $Cap_{q,N}$  is sub-additive, that is  $Cap_{q,N}(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} Cap_{q,N}(A_i)$ .

If  $A$  is Borel measurable, then  $\mathbb{P}(A) \leq Cap_{q,N}(A)^q$  for any  $q > 1$  and  $N \geq 0$ . Therefore any subset of  $\mathbf{W}$  with  $(q, N)$ -capacity zero is null set with respect to the Wiener measure  $\mathbb{P}$ . However there are many probability zero subsets which have positive  $(q, 1)$ -capacity. According to P. Malliavin, a subset  $A \subset \mathbf{W}$  is called *slim*, if  $Cap_{q,N}(A) = 0$  for all  $q \geq 1$  and  $N \in \mathbb{N}$ .

We are interested in the random field  $\mathbf{B}(w) = (1, w^1, w^2)$  on  $\mathbf{W}$ , valued in  $T^2(\mathbb{R}^d)$  parameterized by  $\{(s, t) : s < t\}$ , where

$$\begin{aligned} w_{s,t}^1 &= w_t - w_s, \\ w_{s,t}^2 &= \int_{s < t_1 < t_2 < t} \circ dw_{t_1} \otimes \circ dw_{t_2} \end{aligned}$$

for all  $s < t$ , where  $\circ d$  indicates the Stratonovich integration.

We are now in a position to state an interesting consequence to our main result Theorem 2.3.

**Theorem 2.2** *For any  $p \in (2, 3)$  there is a version of  $\mathbf{B}$  such that  $\{\mathbf{B} \notin G\Omega_p(\mathbb{R}^d)\}$  is slim, that is*

$$Cap_{q,N} \{w \in \mathbf{W} : \mathbf{B}(w) \notin G\Omega_p(\mathbb{R}^d)\} = 0$$

for any  $q > 1$  and  $N \in \mathbb{N}$ .

It is well known and indeed it is very easy to show that there is a version of the stochastic integrals  $B^2$  (defined by (2.1)) so that  $\mathbf{B}$  is well defined on  $\mathbf{W}$  except for a slim subset (i.e. a subset with  $(q, N)$ -capacity zero for all  $q$  and all positive integer  $N$ ), and  $\mathbf{B}$  is a rough path quasi-surely, that is,  $\mathbf{B}(w)$  has finite  $p$ -variation and satisfies Chen's equation for all  $w \in \mathbf{W}$  except for a slim subset. Since  $2 < p < 3$ , such a version of  $\mathbf{B}$  allows us to develop a theory of stochastic differential equations and thus gives quasi-surely defined solutions for all stochastic differential equations with coefficients which are regular enough. However, such a theory will not ensure a convergence theorem such as Theorem 2.4. In fact, we will prove a quasi-sure approximation theorem for the Brownian motion.

For a given natural number  $n$ ,  $k = 0, 1, \dots, 2^n$ ,  $t_n^k = k/2^n$  are the dyadic points in  $[0, 1]$ . For a continuous path  $w \in \mathbf{W}$ ,  $w^{(n)}$  is the polygonal approximation defined by

$$w_t^{(n)} = w_{t_n^{k-1}} + 2^n(t - t_n^{k-1})(w_{t_n^k} - w_{t_n^{k-1}}) \quad \text{for } t \in [t_n^{k-1}, t_n^k] \quad (2.2)$$

for  $k = 1, \dots, 2^n$ . This notation equally applies to the coordinate process  $\{w_t : t \in [0, 1]\}$ .

The idea of approximating Brownian motion by piece-wise smooth sample paths originated from the fundamental research of P. Lévy [33], [34], also see K. Itô and H. P. McKean [26] for the construction of Brownian motion sample paths starting from polygonal paths with vortices modelled by random walks.

For simplicity, if no confusion is possible, we will write  $\mathbf{w}^{(m)}$  for  $L_2(w^{(m)})$ , the enhanced geometric rough path of level two which is associated to  $w^{(m)}$ . Our main result may be stated as the following



**Theorem 2.3** For any  $p \in (2, 3)$

$$Cap_{q,N} \left\{ w \in \mathbf{W} : \sum_{m=1}^{\infty} d_p(\mathbf{w}^{(m)}, \mathbf{w}^{(m+1)}) = \infty \right\} = 0 \quad (2.3)$$

for any  $q \geq 1$  and  $N \in \mathbb{N}$ .

It is obvious that (2.3) implies that for any  $p$  between 2 and 3 there is a subset  $A \subset \mathbf{W}$  such that  $Cap_{q,N}(A) = 0$  for all  $q > 1$ ,  $N \in \mathbb{N}$ , and  $\mathbf{w}^{(m)}$  converges in  $G\Omega_p(\mathbb{R}^d)$  to a limit  $\mathbf{w}$  on  $\mathbf{W} \setminus A$ , which is a modification of  $\mathbf{B}$ .

Putting together with Lyons' universal limit theorem, we obtain immediately the following quasi-sure limit theorem.

**Theorem 2.4** Consider the Stratonovich's type stochastic differential equations on the Wiener space  $(\mathbf{W}, \mathcal{B}(\mathbf{W}), \mathbb{P})$  (so that the coordinate process  $w = (w_t)_{t \geq 0}$  is a standard Brownian motion)

$$dy_t = \sum_{i=1}^d f_i(y_t) \circ dw_t^i + f_0(y_t) dt \quad (2.4)$$

with initial data  $y_0$ ,  $f_i = (f_i^j)$  ( $i = 0, \dots, d$ ;  $j = 1, \dots, N$ ). Suppose  $f_i^j$  and  $f^j$  are in  $C_b^3(\mathbb{R}^N)$ . Suppose for each  $m$ ,  $y^{(m)}$  be the unique solution to the ordinary equation

$$dy_t = \sum_{i=1}^d f_i(y_t) dw_t^{(m),i} + f_0(y_t) dt.$$

Then for any  $p \in (2, 3)$  there is a slim subset  $A \subset \mathbf{W}$  which is independent of (2.4) such that  $\mathbf{y}^{(m)} = L_2(y^{(m)})$  converges to  $\mathbf{y} = (1, y^1, y^2)$  on  $\mathbf{W} \setminus A$ , and  $y_t = y_0 + y_{0,t}^1$  is a version of the unique strong solution to (2.4).

This kind of limit theorems for stochastic differential equations via ordinary differential equations in the context of almost sure sense have been discussed by E. McShane [42], E. Wong and M. Zakai [50], D. Stroock and S.R.S. Varadhan [47], and etc., see Section 7, Chapter VI in N. Ikeda and S. Watanabe [24] for a definite form and for further reference therein. By using Lyons' universal limit theorem, the Wong-Zakai type limit theorem has been extended to other rough differential equation driven by symmetric diffusions in [21], [21], by fractional Brownian motions in [6] and by other Gaussian processes in [35], [12], and A. M. Davie [7], [8] and etc.

In the context of quasi-sure analysis, partial results (i.e. for a solution to a single differential equation or special Wiener functionals) have been obtained by T. Kazumi [28], Z. Huang and J. G. Ren [23], J. G. Ren [44], P. Malliavin and D. Nualart [41], S. Fang [10]. Our result is a natural generalization of the preceding mentioned results, and our negligible set is universal which is independent of the concerned Wiener functionals.

The capacity theory on the infinite dimensional space  $\mathbf{W}$  was first studied by P. Malliavin [39]. In fact, Malliavin introduced the concept of slim sets as negligible sets – those subsets of  $\mathbf{W}$  with  $(q, N)$ -capacity zero for all  $q > 1$  and positive integer  $N$ , by using the Ornstein-Uhlenbeck process on the Wiener space  $\mathbf{W}$ . The current definition of  $Cap_{q,N}$  was gradually

developed through a series of work by I. Shigekawa [46], H. Sugita [48], M. Fukushima [15], [13], [18], [17], [16], H. Kaneko [27], M. Takeda [49], J. G. Ren [45], F. Hirsch and S. Song [22], see P. Malliavin [40] and K. Itô [25] for systematic expositions on slim sets, and M. Fukushima [14], Z.M. Ma and M. Röckner [38] and N. Bouleau and F. Hirsch [3] for the capacity theory defined via analytic or probabilistic potential theory.

Many important almost sure properties which hold for Brownian motion were proved for the corresponding quasi-sure versions by D. Williams (see the article by P. Meyer [43]), M. Fukushima [16], M. Takeda [49], S. Fang [11] etc. S. Kusuoka [29], [30] initiated the study of non-linear analysis on abstract Wiener spaces by using capacity theory.

The starting point in our approach is the capacity version of the Borel Cantelli lemma, that is, if  $\sum_{m=1}^{\infty} \text{Cap}_{q,s}(A_m) < \infty$  then  $\text{Cap}_{q,s}(\overline{\lim}_{m \rightarrow \infty} A_m) = 0$ .

Let  $p \in (2, 3)$  be fixed, and consider

$$A_m = \left\{ w \in \mathbf{W} : d_p(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) > C \left( \frac{1}{2^m} \right)^{\beta} \right\}$$

for some  $\beta > 0$  and constant  $C > 0$ . If we are able to show that

$$\sum_{m=1}^{\infty} \text{Cap}_{q,N}(A_m) < \infty \quad (2.5)$$

then  $\text{Cap}_{q,N}(\overline{\lim}_{m \rightarrow \infty} A_m) = 0$  so that

$$\text{Cap}_{q,N} \left\{ \sum_{m=1}^{\infty} d_p(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) = \infty \right\} \leq \text{Cap}_{q,N}(\overline{\lim}_{m \rightarrow \infty} A_m) = 0$$

which yields Theorem 2.3. Therefore, we would like to estimate

$$\text{Cap}_{q,N} \{ w \in \mathbf{W} : d_p(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) > \lambda \}. \quad (2.6)$$

There are few techniques available to bound the capacity such as (2.6) in contrast to corresponding almost sure statements. In fact the only effective tool to the knowledge of the present authors is the capacity maximal inequality (also called the Tchebycheff inequality, see 1.2.5 on page 92 and 2.2 on page 96 in [40]), which says that, if  $u \in \mathbb{D}_N^q$  and if  $u$  is lower semi-continuous or continuous with respect to the capacity  $\text{Cap}_{q,s}$ , then

$$\text{Cap}_{q,N} \{ w \in \mathbf{W} : u(w) > \lambda \} \leq \frac{C_{q,N}}{\lambda} \|u\|_{q,N} \quad \forall \lambda > 0 \quad (2.7)$$

where  $C_{q,N}$  is a constant depending only on  $d, q$  and  $N$ . This requires to estimate the Sobolev norms of  $u$ . Unfortunately, we are unable to show (and we do not believe it is true) that  $w \rightarrow d_p(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})$  is differentiable in the Malliavin sense. Instead we consider a metric over paths which dominates the  $p$ -variation distance, but differentiable in Malliavin sense and still good enough for Brownian motion.

In what follows,  $p \in (2, 3)$  and  $\gamma > \frac{p}{2} - 1$  be fixed.

If  $\mathbf{w} = (1, w^1, w^2)$  and  $\tilde{\mathbf{w}} = (1, \tilde{w}^1, \tilde{w}^2)$  are two functions on  $\Delta$  taking values in  $T^2(\mathbb{R}^d)$ , we consider

$$\rho_j(\mathbf{w}, \tilde{\mathbf{w}}) = \left( \sum_{n=1}^{\infty} n^{\gamma} \sum_{k=1}^{2^n} \left| w_{t_n^{k-1}, t_n^k}^j - \tilde{w}_{t_n^{k-1}, t_n^k}^j \right|^{\frac{p}{j}} \right)^{\frac{j}{p}} \quad (2.8)$$

where  $j = 1$  or  $2$ . We will use  $\rho_j(\mathbf{w})$  to denote  $\rho_j(\mathbf{w}, \tilde{\mathbf{w}})$  with  $\tilde{\mathbf{w}} = (1, 0, 0)$ .  $\rho_j$  were invented and used in B. Hambly and T. Lyons [21] for constructing the stochastic area processes associated with Brownian motions on the Sierpinski gasket. These functionals were used in M. Ledoux, Z. Qian and T. Zhang [31] to show the large deviation principle for Brownian motion under the topology generated by the  $p$ -variation distance. The following estimates have been contained implicitly in [21] and made explicit in [35] and [31].

**Lemma 2.5** *Suppose  $\gamma > \frac{p}{2} - 1$ . Then there is a positive constant  $C$  depending only on  $\gamma$ ,  $d$  and  $p$  such that*

$$\left( \sup_D \sum_l \left| w_{t_{l-1}, t_l}^1 \right|^p \right)^{\frac{1}{p}} \leq C \rho_1(\mathbf{w}), \quad (2.9)$$

$$\left( \sup_D \sum_l \left| w_{t_{l-1}, t_l}^2 \right|^{\frac{p}{2}} \right)^{\frac{2}{p}} \leq C (\rho_1(\mathbf{w})^2 + \rho_2(\mathbf{w})) \quad (2.10)$$

where  $\sup_D$  takes over finite partitions  $D$  of  $[0, 1]$ , and

$$d_p(\mathbf{w}, \tilde{\mathbf{w}}) \leq C \max \{ \rho_1(\mathbf{w}, \tilde{\mathbf{w}}), \rho_2(\mathbf{w}, \tilde{\mathbf{w}}), \rho_1(\mathbf{w}, \tilde{\mathbf{w}}) (\rho_1(\mathbf{w}) + \rho_1(\tilde{\mathbf{w}})) \}. \quad (2.11)$$

The idea in our approach is to replace  $d_p$  by the right-hand side of (2.11). Therefore, instead of considering the capacity of  $\{d_p(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) > \lambda\}$ , we consider the following

$$\rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) = \left( \sum_{n=1}^{\infty} n^{\gamma} \sum_{k=1}^{2^n} \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{p}{j}} \right)^{\frac{j}{p}}$$

where  $j = 1, 2$  and  $m = 1, 2, \dots$ . Let

$$u_j^{(m)}(w) \equiv \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{j}} = \sum_{n=1}^{\infty} n^{\gamma} \sum_{k=1}^{2^n} \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{p}{j}}.$$

Then, clearly, for each  $N = 1, 2, \dots$ ,

$$u_j^{(m),N}(w) = \sum_{n=1}^N n^{\gamma} \sum_{k=1}^{2^n} \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{p}{j}}$$

is continuous on  $\mathbf{W}$  (which is equipped with the uniform norm over  $[0, 1]$ ), and  $u_j^{(m)}(w) = \sup_N u_j^{(m),N}(w)$ . Therefore  $u_j^{(m)}$  is lower semi-continuous on  $\mathbf{W}$ , and moreover  $u_j^{(m)} \in \mathbb{D}_1^q$  for any  $q \geq 1$ . Therefore we may apply (2.7) to deduce that

$$\text{Cap}_{q,1} \left\{ \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{j}} > \lambda \right\} \leq \frac{C_q}{\lambda} \left\| \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{j}} \right\|_{q,1} \quad (2.12)$$

and similarly

$$\begin{aligned} & Cap_{q,1} \{ \rho_1(\mathbf{w}^{(m)})^p \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p > \lambda \} \\ & \leq \frac{C_q}{\lambda} \| \rho_1(\mathbf{w}^{(m)})^p \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p \|_{q,1} \end{aligned} \quad (2.13)$$

for any  $\lambda > 0$ , where  $C_q > 0$  depends only on  $q$  and  $d$ . In the next section we will establish the necessary estimates to ensure (2.5) for the case that  $N = 1$ .

This approach can not be extended to  $(q, N)$ -capacity case with  $N \geq 2$ , this is because of a simple reason that our dominated distance  $\rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{j}}$  does not belong to  $\mathbb{D}_N^q$  for  $N \geq 2$ . We need new idea to estimate the  $(q, N)$ -capacity for  $N \geq 2$ . The observation to get around this difficulty is that the capacity of  $\{ \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{p/j} > \lambda \}$  is “evenly” distributed over the dyadic partitions, which allows to reduce our task to estimating the capacities of some polynomials of Brownian motion sample paths, which will be explained in the last section of the article, where we conclude the proof of Theorem 2.3.

### 3 Some technical estimates

In this section we establish several technical estimates which will be used in the construction quasi-surely defined geometric rough paths associated with Brownian motion.

We are going to use the following notations. If  $J \subset [0, \infty)$  is a finite interval, then  $1_J$  is the characteristic function of  $J$ , and  $\mathbf{1}_J \in \mathbf{H}$ , which is  $\mathbb{R}^d$ -valued function with the same component  $\int_0^\cdot 1_J(s) ds$ . Hence  $|\mathbf{1}_J|_{\mathbf{H}} = \sqrt{d} \sqrt{|J|}$  where  $|J|$  denotes the length of the interval  $J$ . We will frequently use the following elementary fact: if  $\{J_i : i = 1, \dots, n\}$  is a family of *disjoint* finite intervals, then

$$\left| \sum_{i=1}^n \mathbf{1}_{J_i} \right|_{\mathbf{H}} = \sqrt{d} \sqrt{\sum_{i=1}^n |J_i|}. \quad (3.1)$$

The corresponding fact for the increments of Brownian motion instead of characteristic functions is the context of the following lemma.

**Lemma 3.1** *Then there is a constant  $C > 0$  depending only on  $d$ , such that*

$$\left\| \sum_{i=1}^N \xi_i \otimes \tilde{\xi}_i \right\|_q \leq Cq\sqrt{N} \quad (3.2)$$

for any  $\xi_i, \tilde{\xi}_j$  which are independent valued in  $\mathbb{R}^d$ , with the standard normal  $N(0, 1_{\mathbb{R}^d})$ , for any  $q \geq 1$  and  $N \in \mathbb{N}$ .

This lemma follows from a simple application of the hypercontractivity of the O-U semi-group.

The increment over the interval  $J_n^k \equiv (t_n^{k-1}, t_n^k]$  of  $w \in \mathbf{W}$  is denoted by  $\xi_n^k(w)$  or simply by  $\xi_n^k$  (which denotes the functional  $w \rightarrow \xi_n^k(w)$  as well by abusing the notation), if no confusion may arise. That is

$$\xi_n^k = w_{\frac{k}{2^n}} - w_{\frac{k-1}{2^n}}, \quad k = 1, \dots, 2^n. \quad (3.3)$$

For each  $n$ , since  $\{J_n^k : k = 1, \dots, 2^n\}$  are disjoint,  $\{\xi_n^k : k = 1, \dots, 2^n\}$  are independent, identically distributed with normal distribution  $N(0, 2^{-n}I_{\mathbb{R}^d})$ .

Recall that  $\mathbf{w}^{(n)} = L_2(w^{(n)})$  with first level component  $w^{(n),1}$  and second level  $w^{(n),2}$  respectively, so that

$$\begin{cases} w_{s,t}^{(n),1} &= w_t^{(n)} - w_s^{(n)}, \\ w_{s,t}^{(n),2} &= \int_{s < t_1 < t_2 < t} dw_{t_1}^{(n)} \otimes dw_{t_2}^{(n)}. \end{cases} \quad (3.4)$$

It is easy to see that

$$w_{t_n^{k-1}, t_n^k}^{(m),1} = \begin{cases} \xi_n^k & \text{for } n < m, \\ \frac{2^m}{2^n} \xi_m^{k(n,m)} & \text{for } n \geq m \end{cases} \quad (3.5)$$

where  $k = 1, \dots, 2^n$ , and in the case  $n > m$ ,  $k(n, m)$  is the unique integer  $l$  between 1 and  $2^m$  such that

$$t_m^{l-1} \leq t_n^{k-1} < t_n^k < t_m^l. \quad (3.6)$$

In order to write down some formulas which will be used in what follows, it is better to use Poisson bracket operations  $[\cdot, \cdot]$  and  $\{\cdot, \cdot\}$ , that is, if  $\xi, \eta \in \mathbb{R}^d$ , then

$$[\xi, \eta] = \xi \otimes \eta - \eta \otimes \xi \quad (3.7)$$

and

$$\{\xi, \eta\} = \xi \otimes \eta + \eta \otimes \xi, \quad (3.8)$$

while we will reserve the sharp bracket  $\langle a, b \rangle$  to denote the scalar product in the Euclidean spaces, or in Hilbert space  $\mathbf{H}^{\otimes k}$ .

With these notations, we have (see Section 4.2 in [35] for details)

$$w_{t_n^{k-1}, t_n^k}^{(m),2} = \frac{1}{2} \xi_n^k \otimes \xi_n^k + \frac{1}{2} \sum_{\substack{r,s=2^{m-n}(k-1)+1 \\ r < s}}^{2^{m-n}k} [\xi_m^r, \xi_m^s] \quad (3.9)$$

for  $n < m$ , so that, if  $n < m$ , then

$$w_{t_n^{k-1}, t_n^k}^{(m+1),2} - w_{t_n^{k-1}, t_n^k}^{(m),2} = \frac{1}{2} \sum_{r=2^{m-n}(k-1)+1}^{2^{m-n}k} [\xi_{m+1}^{2r-1}, \xi_{m+1}^{2r}]. \quad (3.10)$$

If  $n \geq m$ , then

$$w_{t_n^{k-1}, t_n^k}^{(m),2} = \frac{1}{2} \frac{2^{2m}}{2^{2n}} \xi_m^{k(n,m)} \otimes \xi_m^{k(n,m)}. \quad (3.11)$$

### 3.1 Moment estimates under Sobolev norms

In this part, let  $p \in (2, 3)$  is a constant,  $d$  is the dimension,  $n, m \in \mathbb{N}$ . We wish to develop several moment estimates for  $w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j}$  (for  $j = 1, 2$ ).

**Lemma 3.2** *There is a constant  $C$  depending only on  $d$  such that*

$$\left\| w_{t_n^{k-1}, t_n^k}^{(m),j} \right\|_q \leq \begin{cases} C \left( \sqrt{q} \frac{1}{\sqrt{2^n}} \right)^j & \text{for } n < m, \\ C \left( \sqrt{q} \frac{\sqrt{2^m}}{2^n} \right)^j & \text{for } n \geq m \end{cases} \quad (3.12)$$

where  $j = 1, 2$ ,

$$\left\| w_{t_n^{k-1}, t_n^k}^{(m+1),1} - w_{t_n^{k-1}, t_n^k}^{(m),1} \right\|_q \leq \begin{cases} 0, & \text{if } n \leq m, \\ C \sqrt{q} \sqrt{\frac{2^m}{2^{2n}}} & \text{if } n \geq m, \end{cases} \quad (3.13)$$

and

$$\left\| w_{t_n^{k-1}, t_n^k}^{(m+1),2} - w_{t_n^{k-1}, t_n^k}^{(m),2} \right\|_q \leq \begin{cases} C q \sqrt{\frac{1}{2^{m+n}}} & \text{if } n \leq m, \\ C q \frac{2^m}{2^{2n}} & \text{if } n \geq m. \end{cases} \quad (3.14)$$

for any  $q \geq 1$ .

**Proof.** For simplicity, let  $Y_j = w_{t_n^{k-1}, t_n^k}^{(m),j}$  and  $X_j = w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j}$ . For  $j = 1$ , (3.12) follows from the fact that

$$\|Y_1\|_q = \begin{cases} \frac{1}{\sqrt{2^n}} \|\xi\|_q & \text{for } n \leq m, \\ \frac{\sqrt{2^m}}{2^n} \|\xi\|_q & \text{for } n > m \end{cases} \quad (3.15)$$

where  $\xi \sim N(0, 1_{\mathbb{R}^d})$ , and the fact that  $\|\xi\|_q \leq C\sqrt{q}$  for some constant  $C$  depending only on  $d$ . For  $j = 2$  and  $n > m$ , (3.12) follows from (3.11) directly. Consider the case that  $n < m$ . According to (3.9), we need to estimate

$$I_q = \left\| \sum_{\substack{r, s=2^{m-n}(k-1)+1 \\ r < s}}^{2^{m-n}k} \xi_m^r \otimes \xi_m^s \right\|_q = \left\| \sum_{s=2^{m-n}(k-1)+1}^{2^{m-n}k} \sum_{r=2^{m-n}(k-1)+1}^{s-1} \xi_m^r \otimes \xi_m^s \right\|_q.$$

Since  $\xi_m^r \otimes \xi_m^s$  belong to the second chaos component for whatever  $m$ , so that  $I_q \leq C_1 q I_2$  for some constants  $C_1$  and  $C_2$  depending only on  $d$ , but independent of  $m$  or  $n$ . Therefore we may assume that  $q = 2$ . Furthermore, for simplicity, set  $\eta_s = \sum_{r=2^{m-n}(k-1)+1}^{s-1} \xi_m^r$ . Then  $\xi_s$  has a normal distribution with mean zero and

$$\text{var}(\eta_s) = \frac{1}{2^m} (s - 2^{m-n}(k-1)).$$

Thus

$$\begin{aligned} \left\| \sum_{s=2^{m-n}(k-1)+1}^{2^{m-n}k} \eta_s \otimes \xi_m^s \right\|_2^2 &= \sum_{i,j=1}^d \mathbb{E} \left( \sum_{s=2^{m-n}(k-1)+1}^{2^{m-n}k} \eta_s^i \xi_m^{s,j} \right)^2 \\ &= \sum_{i,j=1}^d \sum_{s=2^{m-n}(k-1)+1}^{2^{m-n}k} \mathbb{E} (\eta_s^i \xi_m^{s,j})^2 \\ &\quad + 2 \sum_{i,j=1}^d \sum_{\substack{s=2^{m-n}(k-1)+1 \\ s < r}}^{2^{m-n}k} \mathbb{E} (\eta_s^i \xi_m^{s,j} \eta_r^i \xi_m^{r,j}). \end{aligned}$$

The last term has contribution zero. This is because for  $r > s$ ,  $\xi_m^{r,j}$  is independent of  $\eta_s^i \xi_m^{s,j} \eta_r^i$ , so that

$$\mathbb{E}(\eta_s^i \xi_m^{s,j} \eta_r^i \xi_m^{r,j}) = \mathbb{E}(\eta_s^i \xi_m^{s,j} \eta_r^i) \mathbb{E} \xi_m^{r,j} = 0 \quad \text{for } s < r.$$

Hence

$$\begin{aligned} \left\| \sum_{s=2^{m-n}(k-1)+1}^{2^{m-n}k} \eta_s \otimes \xi_m^s \right\|_2^2 &= \sum_{i,j=1}^d \sum_{s=2^{m-n}(k-1)+1}^{2^{m-n}k} \mathbb{E}(\eta_s^i \xi_m^{s,j})^2 \\ &= \sum_{i,j=1}^d \sum_{s=2^{m-n}(k-1)+1}^{2^{m-n}k} \mathbb{E}(\xi_m^{s,j})^2 \mathbb{E}(\eta_s^i)^2 \\ &= d^2 \sum_{s=2^{m-n}(k-1)+1}^{2^{m-n}k} \frac{1}{2^m} (s - 2^{m-n}(k-1)) \frac{1}{2^m} \\ &= \frac{1}{2^{2m}} d^2 \sum_{s=1}^{2^{m-n}} s \leq C \frac{1}{2^{2n}} \end{aligned}$$

and therefore

$$\left\| \sum_{\substack{r,s=2^{m-n}(k-1)+1 \\ r < s}}^{2^{m-n}k} \xi_m^r \otimes \xi_m^s \right\|_q \leq C_2 q \frac{1}{2^n}$$

for some constant  $C$  depending only on  $d$ . By (3.4) and the preceding estimate we have

$$\begin{aligned} \|Y_2\|_q &\leq \frac{1}{2} \|\xi_n^k \otimes \xi_n^k\|_q + \frac{1}{2} \left\| \sum_{\substack{r,s=2^{m-n}(k-1)+1 \\ r < s}}^{2^{m-n}k} [\xi_m^r, \xi_m^s] \right\|_q \\ &\leq \frac{1}{2} \|\xi_n^k \otimes \xi_n^k\|_q + \left\| \sum_{\substack{r,s=2^{m-n}(k-1)+1 \\ r < s}}^{2^{m-n}k} \xi_m^r \otimes \xi_m^s \right\|_q \\ &\leq C \frac{q}{2^n}. \end{aligned}$$

Remain to show (3.14) for the case  $n \leq m$ . Indeed, by (3.10) and (3.2)

$$\begin{aligned} \left\| w_{t_n^{k-1}, t_n^k}^{(m+1),2} - w_{t_n^{k-1}, t_n^k}^{(m),2} \right\|_q &= \frac{1}{2} \left\| \sum_{r=2^{m-n}(k-1)+1}^{2^{m-n}k} [\xi_{m+1}^{2r-1}, \xi_{m+1}^{2r}] \right\|_q \\ &\leq Cq \sqrt{\frac{1}{2^{m+n}}} \end{aligned}$$

which completes the proof of the lemma. ■

**Lemma 3.3** *There is a constant  $C$  depending only on  $d$  such that If  $n \geq m$ , then  $Y_1 = \frac{2^m}{2^n} \xi_m^{k(n,m)}$ ,  $DY_1 = \frac{2^m}{2^n} \mathbf{1}_{J_m^{k(n,m)}}$  If  $n < m$ , then  $Y_1 = \xi_n^k$  so that  $DY_1 = \mathbf{1}_{J_n^k}$*

$$\left| Dw_{t_n^{k-1}, t_n^k}^{(m),1} \right|_{\mathbf{H}} \leq \begin{cases} C \sqrt{\frac{1}{2^n}} & \text{if } n \leq m, \\ C \frac{2^m}{2^n} \sqrt{\frac{1}{2^m}} & \text{if } n > m, \end{cases} \quad (3.16)$$

$$\left| D \left( w_{t_n^{k-1}, t_n^k}^{(m+1),1} - w_{t_n^{k-1}, t_n^k}^{(m),1} \right) \right|_{\mathbf{H}} \leq \begin{cases} 0 & \text{if } n \leq m, \\ C \sqrt{\frac{2^m}{2^n}} \sqrt{\frac{1}{2^n}} & \text{if } n > m \end{cases} \quad (3.17)$$

and

$$\left\| D^a \left( w_{t_n^{k-1}, t_n^k}^{(m+1),2} - w_{t_n^{k-1}, t_n^k}^{(m),2} \right) \right\|_q \leq \begin{cases} C \sqrt{q}^{2-a} \sqrt{\frac{1}{2^{m+n}}} & \text{if } n \leq m, \\ C \sqrt{q}^{2-a} \sqrt{\frac{2^{3m}}{2^{3n}}} \sqrt{\frac{1}{2^{n+m}}} & \text{if } n > m \end{cases} \quad (3.18)$$

for any  $m, n \in \mathbb{N}$ , and  $k = 1, \dots, 2^n$ ,  $a = 1, 2$ .

**Proof.** (3.16) is easy to see. Let  $X_j = w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j}$  for simplicity. If  $n \leq m$  then  $X_1 = 0$ , and if  $n \geq m$ , then

$$DX_1 = \frac{2^{m+1}}{2^n} \mathbf{1}_{J_{m+1}^{k(n,m+1)}} - \frac{2^m}{2^n} \mathbf{1}_{J_m^{k(n,m)}}$$

so that

$$|DX_1|_{\mathbf{H}} = C \sqrt{\frac{2^m}{2^n}} \sqrt{\frac{1}{2^n}}$$

where  $C$  is a constant depending only on  $d$ . Next we consider the Lévy area  $X_2$ . If  $n < m$

$$X_2 = \frac{1}{2} \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} [\xi_{m+1}^{2l-1}, \xi_{m+1}^{2l}].$$

so that

$$DX_2 = \frac{1}{2} \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} \left( [\mathbf{1}_{J_{m+1}^{2l-1}}, \xi_{m+1}^{2l}] + [\xi_{m+1}^{2l-1}, \mathbf{1}_{J_{m+1}^{2l}}] \right)$$

and

$$D^2X_2 = \frac{1}{2} \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} \left( [\mathbf{1}_{J_{m+1}^{2l-1}}, \mathbf{1}_{J_{m+1}^{2l}}] - [\mathbf{1}_{J_{m+1}^{2l}}, \mathbf{1}_{J_{m+1}^{2l-1}}] \right)$$

where

$$[\mathbf{1}_{J_{m+1}^{2l-1}}, \mathbf{1}_{J_{m+1}^{2l}}](t_1, t_2) = [\mathbf{1}_{J_{m+1}^{2l-1}}(t_1), \mathbf{1}_{J_{m+1}^{2l}}(t_2)].$$



Since the intervals  $J_{m+1}^i$  are disjoint, so that

$$\begin{aligned}
|DX_2|_{\mathbf{H}}^2 &= \int_0^\infty \frac{1}{4} \sum_{i,j} \left( \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} \left( [1_{J_{m+1}^{2l-1}}, \xi_{m+1}^{2l}]^{ij} + [\xi_{m+1}^{2l-1}, 1_{J_{m+1}^{2l}}]^{ij} \right) \right)^2 \\
&= \int_0^\infty \frac{1}{4} \sum_{i,j} \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} \left( [1_{J_{m+1}^{2l-1}}, \xi_{m+1}^{2l}]^{ij} + [\xi_{m+1}^{2l-1}, 1_{J_{m+1}^{2l}}]^{ij} \right)^2 \\
&\leq \frac{1}{2} \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} \int_0^\infty \left( |[1_{J_{m+1}^{2l-1}}, \xi_{m+1}^{2l}]|^2 + |[\xi_{m+1}^{2l-1}, 1_{J_{m+1}^{2l}}]|^2 \right) \\
&\leq \frac{1}{2^{m+1}} \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} (|\xi_{m+1}^{2l}|^2 + |\xi_{m+1}^{2l-1}|^2)
\end{aligned}$$

that is

$$|DX_2|_{\mathbf{H}} \leq \sqrt{\frac{1}{2^{m+1}}} \sqrt{\sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} (|\xi_{m+1}^{2l}|^2 + |\xi_{m+1}^{2l-1}|^2)}. \quad (3.19)$$

Similarly,

$$\begin{aligned}
|D^2X_2|_{\mathbf{H}^{\otimes 2}}^2 &= \int_0^\infty \int_0^\infty \frac{1}{4} \sum_{i,j} \left( \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} \left( [1_{J_{m+1}^{2l-1}}, 1_{J_{m+1}^{2l}}] - [1_{J_{m+1}^{2l}}, 1_{J_{m+1}^{2l-1}}] \right)^{ij} \right)^2 \\
&= \int_0^\infty \int_0^\infty \frac{1}{4} \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} \sum_{i,j} |[1_{J_{m+1}^{2l-1}}, 1_{J_{m+1}^{2l}}] - [1_{J_{m+1}^{2l}}, 1_{J_{m+1}^{2l-1}}]|^2 \\
&\leq \int_0^\infty \int_0^\infty \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} |[1_{J_{m+1}^{2l-1}}, 1_{J_{m+1}^{2l}}]| \\
&= C 2^{m-n} \frac{1}{2^{2(m+1)}}
\end{aligned}$$

where  $C$  depends only on  $d$ , so that

$$|D^2X_2|_{\mathbf{H}^{\otimes 2}} \leq C \sqrt{\frac{1}{2^{m+n}}}. \quad (3.20)$$

Hence, for  $q \geq 2$  we have

$$\begin{aligned}
\|DX_2\|_q &\leq \sqrt{\frac{1}{2^{m+1}}} \sqrt{\left\| \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} (|\xi_{m+1}^{2l}|^2 + |\xi_{m+1}^{2l-1}|^2) \right\|_{\frac{q}{2}}} \\
&\leq \sqrt{\frac{1}{2^{m+1}}} \sqrt{\sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} (\|\xi_{m+1}^{2l}\|_q^2 + \|\xi_{m+1}^{2l-1}\|_q^2)} \\
&\leq C \sqrt{q} \sqrt{\frac{1}{2^{m+1}}} \sqrt{2^{m-n} \frac{1}{2^m}} \leq C \sqrt{q} \sqrt{\frac{1}{2^{m+n}}}
\end{aligned}$$

and

$$\|D^2 X_2\|_q \leq C \frac{1}{\sqrt{2^{m+n}}} \quad \forall n \leq m. \quad (3.21)$$

If  $n \geq m$  then

$$\begin{aligned} DX_2 &= \frac{1}{2} \frac{2^{2m}}{2^{2n}} \left( 4\{\xi_{m+1}^{k(n,m+1)}, \mathbf{1}_{J_{m+1}^{k(n,m+1)}}\} - \{\xi_m^{k(n,m)}, \mathbf{1}_{J_m^{k(n,m)}}\} \right), \\ D^2 X_2 &= \frac{1}{2} \frac{2^{2m}}{2^{2n}} \left( 4\{\mathbf{1}_{J_{m+1}^{k(n,m+1)}}, \mathbf{1}_{J_{m+1}^{k(n,m+1)}}\} - \{\mathbf{1}_{J_m^{k(n,m)}}, \mathbf{1}_{J_m^{k(n,m)}}\} \right) \end{aligned}$$

so that

$$|DX_2|_{\mathbf{H}} \leq C \sqrt{\frac{2^{3m}}{2^{3n}}} \sqrt{\frac{1}{2^n}} \left( |\xi_{m+1}^{k(n,m+1)}| + |\xi_m^{k(n,m)}| \right), \quad |D^2 X_2|_{\mathbf{H}} \leq C \frac{2^{2m}}{2^{2n}} \frac{1}{2^m}$$

for a constant  $C$  depending only on  $d$ . After taking  $q$ -norm, we obtain (3.18). ■

**Lemma 3.4** *Let  $\tilde{N} > 0$ . Consider the following functionals  $f_{m,n,k}^j$  on the Wiener space  $\mathbf{W}$  defined by*

$$f_{m,n,k}^j(w) = \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{2\tilde{N}} \quad \text{for } w \in \mathbf{W}. \quad (3.22)$$

*If  $2\tilde{N} > 1$  then there is a constant  $C$  depending only on  $d$  and  $\tilde{N}$  such that*

$$\|Df_{m,n,k}^1\|_q \leq C q^{\tilde{N}} \left( \frac{2^m}{2^{2n}} \right)^{\tilde{N}} \quad \text{for } n > m \quad (3.23)$$

and

$$\|Df_{m,n,k}^2\|_q \leq \begin{cases} C \sqrt{q}^{4\tilde{N}} \left( \frac{1}{2^{m+n}} \right)^{\tilde{N}} & \text{if } n \leq m, \\ C \sqrt{q}^{4\tilde{N}} \left( \frac{2^m}{2^{2n}} \right)^{2\tilde{N}} & \text{if } n > m. \end{cases} \quad (3.24)$$

for any  $n, m$  and  $k = 1, \dots, 2^n$ .

**Proof.** Let  $X_j = w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j}$  ( $j = 1, 2$ ) as in the proof of the previous lemma. Since  $2\tilde{N} > 1$  so that  $f_{m,n,k}^j$  are differentiable in the sense of Malliavin calculus, and by chain rule

$$Df_{m,n,k}^j = 2\tilde{N} |X_j|^{2(\tilde{N}-1)} \langle X_j, DX_j \rangle$$

so that

$$|Df_{m,n,k}^j|_{\mathbf{H}} \leq 2\tilde{N} |X_j|^{2\tilde{N}-1} |DX_j|_{\mathbf{H}}.$$

Thus, choose  $\alpha > 1$  such that  $\beta = \alpha q(2\tilde{N} - 1) > 1$ . Then, by using Hölder's inequality

$$\|Df_{m,n,k}^j\|_q \leq 2\tilde{N} \|X_j\|_{\beta}^{2\tilde{N}-1} \|DX_j\|_{q\alpha'} \quad (3.25)$$

where  $\frac{1}{\alpha} + \frac{1}{\alpha'} = 1$ . Using (3.13) and (3.17) to obtain

$$\|Df_{m,n,k}^1\|_q \leq C \left( \sqrt{q} \sqrt{\frac{2^m}{2^n}} \sqrt{\frac{1}{2^n}} \right)^{2\tilde{N}}$$

and

$$\|Df_{m,n,k}^2\|_q \leq \begin{cases} C \left( \sqrt{q}^2 \sqrt{\frac{1}{2^{m+n}}} \right)^{2\tilde{N}-1} \left( \sqrt{q} \sqrt{\frac{1}{2^{m+n}}} \right) & \text{if } n \leq m, \\ C \left( \sqrt{q}^2 \frac{2^m}{2^{2n}} \right)^{2\tilde{N}-1} \left( \sqrt{q} \sqrt{\frac{2^{3m}}{2^{3n}}} \sqrt{\frac{1}{2^{n+m}}} \right) & \text{if } n > m \end{cases}$$

thus (3.23) and (3.24) follow immediately. ■

### 3.2 Estimating capacities

**Lemma 3.5** *There is a constant  $C$  depending only on  $p, \gamma$  and  $d$  such that*

$$\left\| \rho_j(\mathbf{w}^{(m)})^{\frac{p}{j}} \right\|_q \leq C q^{\frac{p}{2}} \quad \forall m \in \mathbb{N}, q \geq 1 \quad (3.26)$$

where  $j = 1, 2$ .

**Proof.** (3.26) follows from the following inequality

$$\left\| \rho_j(\mathbf{w}^{(m)})^{\frac{p}{j}} \right\|_q \leq \sum_{n=1}^{\infty} n^{\gamma} \sum_{k=1}^{2^n} \left\| w_{t_n^{k-1}, t_n^k}^{(m),j} \right\|_{pq}^p$$

and Lemma 3.2 applying to the  $L^{pq}$ -norm. ■

**Lemma 3.6** *There is a constant  $C$  depending only on  $p, \gamma$  and  $d$  such that*

$$\left\| \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{j}} \right\|_q \leq C q^{\frac{p}{2}} \left( \frac{1}{2^m} \right)^{\frac{p-2}{4}} \quad (3.27)$$

for all  $m \in \mathbb{N}$ , and  $j = 1, 2$  and  $q > 1$ .

**Proof.** thus, together with (3.12), there is a constant depending only on  $d$  such that  
Therefore, by applying these estimates to  $L^{pq}$  norms we deduce that

$$\begin{aligned} \left\| \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p \right\|_q &\leq C \sum_{n>m}^{\infty} n^{\gamma} \sum_{k=1}^{2^n} \left\| w_{t_n^{k-1}, t_n^k}^{(m+1),1} - w_{t_n^{k-1}, t_n^k}^{(m),1} \right\|_{qp}^p \\ &\leq C \sum_{n>m}^{\infty} n^{\gamma} 2^n \left( \sqrt{q} \frac{\sqrt{2^m}}{2^n} \right)^p \\ &\leq C q^{\frac{p}{2}} \sum_{n>m}^{\infty} n^{\gamma} \left( \frac{1}{2^n} \right)^{\frac{p-2}{2}} \\ &\leq C q^{\frac{p}{2}} \left( \frac{1}{2^m} \right)^{\frac{p-2}{4}}. \end{aligned}$$

Similarly

$$\begin{aligned}
\left\| \rho_2(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{2}} \right\|_q &\leq \sum_{n=1}^{\infty} n^{\gamma} \sum_{k=1}^{2^n} \left\| w_{t_n^{k-1}, t_n^k}^{(m+1), 2} - w_{t_n^{k-1}, t_n^k}^{(m), j} \right\|_{q^{\frac{p}{2}}}^{\frac{p}{2}} \\
&\leq C q^{\frac{p}{2}} \left( \sum_{n \leq m} n^{\gamma} 2^n \sqrt{\frac{1}{2^{m+n}}}^{\frac{p}{2}} + \sum_{n > m} n^{\gamma} 2^n \frac{2^m}{2^n} \left( \frac{1}{2^n} \right)^{\frac{p}{2}} \right) \\
&\leq C q^{\frac{p}{2}} \sum_{n \leq m} n^{\gamma} \sqrt{\frac{2^n}{2^m}} \left( \frac{1}{2^m} \right)^{\frac{p-2}{4}} \left( \frac{1}{2^n} \right)^{\frac{p-2}{4}} + \sum_{n > m} n^{\gamma} \left( \frac{1}{2^n} \right)^{\frac{p}{2}-1} \\
&\leq C q^{\frac{p}{2}} \left( \frac{1}{2^m} \right)^{\frac{p-2}{4}}.
\end{aligned}$$

■

**Lemma 3.7** *There is a constant  $C$  depending only on  $p, \gamma$  and  $d$  such that*

$$\left\| \rho_1(\mathbf{w}^{(m)})^p \right\|_{q,1} \leq C q^{\frac{p}{2}} \quad (3.28)$$

for any  $q \geq 1$  and  $m \in \mathbb{N}$ .

**Proof.** Let  $Y_1 = w_{t_n^{k-1}, t_n^k}^{(m), 1}$ . Then, by (3.5),  $DY_1 = \mathbf{1}_{J_n^k}$  (if  $n < m$ ) or  $\frac{2^m}{2^n} \mathbf{1}_{J_m^{k(n,m)}}$  for  $n \geq m$ , so that

$$\begin{aligned}
\|D|Y_1|^p\|_q &= p \|Y_1\|_{q(p-1)}^{p-1} |DY_1|_{\mathbf{H}} \\
&= \begin{cases} \left( \frac{1}{\sqrt{2^n}} \|\xi\|_{(p-1)q} \right)^{p-1} \sqrt{\frac{1}{2^n}} & \text{for } n < m \\ \left( \frac{\sqrt{2^m}}{2^n} \|\xi\|_{(p-1)q} \right)^{p-1} \frac{2^m}{2^n} \sqrt{\frac{1}{2^m}} & \text{for } n \geq m \end{cases}
\end{aligned}$$

where  $\xi \sim N(0, 1_{\mathbb{R}^d})$ . Hence

$$\begin{aligned}
\|D\rho_1(\mathbf{w}^{(m)})^p\|_{\mathbf{H}} &\leq \sum_{n=1}^{\infty} n^{\gamma} \sum_{k=1}^{2^n} \left\| D \left| w_{t_n^{k-1}, t_n^k}^{(m), 1} \right|^p \right\|_q \\
&\leq C q^{\frac{p-1}{2}} \sum_{n=1}^{\infty} n^{\gamma} \left( \frac{1}{2^n} \right)^{\frac{p}{2}-1}.
\end{aligned}$$

■

**Lemma 3.8** *There is a constant  $C$  depending only on  $p, \gamma$  and  $d$  such that*

$$\left\| \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{j}} \right\|_{q,1} \leq C q^{\frac{p}{2}} \left( \frac{1}{2^m} \right)^{\frac{p-2}{4}} \quad (3.29)$$

for any  $m, q \geq 1$  and  $j = 1, 2$ .

**Proof.** Let  $u_{m,n,k}^j(w) = \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{p}{j}}$ . According to (3.13), (3.14), (3.23) and (3.24) with  $2\tilde{N} = \frac{p}{j}$ , we may deduce that

$$\|u_{m,n,k}^1\|_{q,1} \leq Cq^{\frac{p}{2}} \left( \frac{2^m}{2^{2n}} \right)^{\frac{p}{2}} \text{ for } n > m$$

and

$$\|u_{m,n,k}^2\|_{q,1} \leq \begin{cases} Cq^{\frac{p}{2}} \left( \frac{1}{2^{m+n}} \right)^{\frac{p}{4}} & \text{if } n \leq m, \\ Cq^{\frac{p}{2}} \left( \frac{2^m}{2^{2n}} \right)^{\frac{p}{2}} & \text{if } n > m. \end{cases}$$

On the other hand, by triangle inequality

$$\left\| \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{j}} \right\|_{q,1} \leq \sum_{n=1}^{\infty} n^{\gamma} \sum_{k=1}^{2^n} \|u_{m,n,k}^j\|_{q,1}.$$

It follows thus that

$$\begin{aligned} \left\| \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p \right\|_{q,1} &\leq Cq^{\frac{p}{2}} \sum_{n>m}^{\infty} n^{\gamma} 2^n \left( \frac{1}{2^n} \right)^{\frac{p-2}{2}} \\ &\leq Cq^{\frac{p}{2}} \left( \frac{1}{2^m} \right)^{\frac{p-2}{4}} \sum_{n>m}^{\infty} n^{\gamma} \left( \frac{1}{2^n} \right)^{\frac{p-2}{4}} \\ &\leq Cq^{\frac{p}{2}} \left( \frac{1}{2^m} \right)^{\frac{p-2}{4}}. \end{aligned}$$

and

$$\begin{aligned} \left\| D\rho_2(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{2}} \right\|_q &\leq Cq^{\frac{p}{2}} \left( \sum_{n \leq m}^{\infty} n^{\gamma} 2^n \left( \frac{1}{2^{m+n}} \right)^{\frac{p}{4}} + \sum_{n > m}^{\infty} n^{\gamma} 2^n \left( \frac{2^m}{2^{2n}} \right)^{\frac{p}{2}} \right) \\ &\leq Cq^{\frac{p}{2}} \left( \frac{1}{2^m} \right)^{\frac{p-2}{4}} \sum_{n=1}^{\infty} n^{\gamma} \left( \frac{1}{2^n} \right)^{\frac{p-2}{4}}. \end{aligned}$$

where constants  $C$  may be different from line to line but only depend on  $p, \gamma$  and  $d$ . ■

Finally we need an  $L^q$ -estimate for the malliavin derivative of  $\rho_1(\mathbf{w}^{(m)})^p \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p$ .

**Lemma 3.9** *There is a constant  $C$  depending only on  $p, \gamma$  and  $d$  such that*

$$\left\| \rho_1(\mathbf{w}^{(m)})^p \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p \right\|_{q,1} \leq Cq^{\frac{p}{2}} \left( \frac{1}{2^m} \right)^{\frac{p-2}{8}} \quad (3.30)$$

for any  $m \in \mathbb{N}$ .

**Proof.** Firstly, by Cauchy-Schwartz's inequality

$$\begin{aligned}
& \left\| \rho_1(\mathbf{w}^{(m)})^p \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p \right\|_q \\
& \leq \sqrt{\left\| \rho_1(\mathbf{w}^{(m)})^p \right\|_{2q}} \sqrt{\left\| \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p \right\|_{2q}} \\
& \leq Cq^{\frac{p}{2}} \left( \frac{1}{2^m} \right)^{\frac{p-2}{8}}.
\end{aligned}$$

Similarly, by chain rule, and Cauchy-Schwartz inequality

$$\begin{aligned}
& \left\| D \left( \rho_1(\mathbf{w}^{(m)})^p \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p \right) \right\|_q \\
& \leq \left\| \rho_1(\mathbf{w}^{(m)})^p D \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p \right\|_q \\
& \quad + \left\| \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p D \rho_1(\mathbf{w}^{(m)})^p \right\|_q \\
& \leq \sqrt{\left\| \rho_1(\mathbf{w}^{(m)})^p \right\|_{2q}} \sqrt{\left\| D \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p \right\|_{2q}} \\
& \quad + \sqrt{\left\| D \rho_1(\mathbf{w}^{(m)})^p \right\|_{2q}} \sqrt{\left\| \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p \right\|_{2q}},
\end{aligned}$$

together with (3.26), (3.28) and (3.29) we obtain (3.30). ■

Recall that  $p \in (2, 3)$  and  $\gamma > \frac{p}{2} - 1$ , and  $C_{p,\gamma}$  the constant appearing in (2.5).

**Theorem 3.10** Suppose  $\beta \in (0, \frac{p-2}{8p})$ , then

$$\sum_m Cap_{q,1} \left\{ d_p(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) > C_{p,\gamma} \left( \frac{1}{2^m} \right)^\beta \right\} < \infty. \quad (3.31)$$

**Proof.** By using our basic estimate (3.29)

$$\begin{aligned}
Cap_{q,1} \left\{ \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) > \lambda^{\frac{1}{p}} \right\} &= Cap_{q,1} \left\{ \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{j}} > \lambda \right\} \\
&\leq \frac{1}{\lambda} \left\| \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{j}} \right\|_{q,1} \\
&\leq \frac{C}{\lambda} \left( \frac{1}{2^m} \right)^{\frac{p-2}{4}}
\end{aligned}$$

where  $C$  is a constant depending only on  $d$  and  $p$ . Choose  $\lambda$  such that

$$\lambda^{j/p} = \left( \frac{1}{2^m} \right)^\beta$$

to obtain

$$Cap_{q,1} \left\{ \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) > \left( \frac{1}{2^m} \right)^\beta \right\} \leq C \left( \frac{1}{2^m} \right)^{\frac{p-2}{4} - \frac{p\beta}{j}}.$$

Since  $\frac{p-2}{4} - \frac{p\beta}{j} > 0$  so that

$$\sum_m Cap_{q,1} \left\{ |\rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})| > \left( \frac{1}{2^m} \right)^\beta \right\} < \infty.$$

Similarly

$$\begin{aligned}
& Cap_{q,1} \left( |\rho_1(\mathbf{w}^{(m)})\rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})| > \lambda^{1/p} \right) \\
& \leq \frac{1}{\lambda} \left\| \rho_1(\mathbf{w}^{(m)})^p \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p \right\|_{q,1} \\
& \leq \frac{C}{\lambda} \left( \frac{1}{2^m} \right)^{\frac{p-2}{8}}
\end{aligned}$$

so that

$$Cap_{q,1} \left\{ \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) > \left( \frac{1}{2^m} \right)^\beta \right\} \leq C \left( \frac{1}{2^m} \right)^{\frac{p-2}{8} - p\beta}.$$

Now (3.31) follows from (2.5). ■

**Corollary 3.11** *Suppose  $p \in (2, 3)$ , then*

$$Cap_{q,1} \left\{ \sum_{m=1}^{\infty} d_p(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) = \infty \right\} = 0, \quad \forall q \geq 1.$$

We have thus proved (2.3) for  $N = 1$ .

## 4 The proof of the quasi-sure convergence

Guided by the estimates we have obtained in the previous section, we wish to show that for every pair  $q \geq 1$  and  $N \in \mathbb{N}$

$$Cap_{q,N} \left\{ \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{j}} > C \left( \frac{1}{2^m} \right)^\beta \right\} \leq C' \left( \frac{1}{2^m} \right)^\varepsilon$$

for some choices of  $\beta > 0$  and  $\varepsilon > 0$ , where  $C$  and  $C'$  are two constants independent of  $m$ .

Therefore we are interested in the capacity

$$I_j(m) = Cap_{q,N} \left\{ \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{j}} > \lambda \right\}$$

( $j = 1, 2$ ).

Since

$$\rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{j}} = \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{p}{j}}$$

so that, for every  $\theta > 0$

$$\left\{ \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^{\frac{p}{j}} > \lambda \right\} \subset \bigcup_{n=1}^{\infty} \left\{ \sum_{k=1}^{2^n} \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{p}{j}} > C_\theta \lambda \left( \frac{1}{2^n} \right)^\theta \right\} \quad (4.1)$$

where

$$C_\theta = \frac{1}{\sum_{n=1}^{\infty} n^\gamma \left(\frac{1}{2^n}\right)^\theta}.$$

Therefore

$$\begin{aligned} I_j(m) &\leq \sum_{n=1}^{\infty} Cap_{q,N} \left\{ \sum_{k=1}^{2^n} \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{p}{j}} > \lambda C_\theta \left(\frac{1}{2^n}\right)^\theta \right\} \\ &\leq \sum_{n=1}^{\infty} \sum_{k=1}^{2^n} Cap_{q,N} \left\{ \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{p}{j}} > \lambda C_\theta \left(\frac{1}{2^n}\right)^{\theta+1} \right\}. \end{aligned} \quad (4.2)$$

On the other hand, for any  $\tilde{N} > 0$  we have

$$\begin{aligned} &Cap_{q,N} \left\{ \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{p}{j}} > \lambda C_\theta \left(\frac{1}{2^n}\right)^{\theta+1} \right\} \\ &= Cap_{q,N} \left\{ \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right| > \lambda^{\frac{j}{p}} C_\theta^{\frac{j}{p}} \left(\frac{1}{2^n}\right)^{\frac{j}{p}(\theta+1)} \right\} \\ &= Cap_{q,N} \left\{ f_{m,n,k}^j > \left[ \lambda^{\frac{j}{p}} C_\theta^{\frac{j}{p}} \left(\frac{1}{2^n}\right)^{\frac{j}{p}(\theta+1)} \right]^{2\tilde{N}} \right\} \end{aligned}$$

where

$$f_{m,n,k}^j(w) = \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{2\tilde{N}} \quad \text{for } w \in \mathbf{W}.$$

If  $\tilde{N}$  is a natural number, then  $f_{m,n,k}^j$  are polynomials of the Brownian motion paths, so are smooth functionals on the Wiener space  $\mathbf{W}$  in Malliavin's sense. This latter fact allows us to apply the capacity maximal inequality to bound the preceding capacity. Namely, for each pair  $q \geq 1$  and  $N \in \mathbb{N}$ , we have

$$Cap_{q,N} \left\{ \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{p}{j}} > \lambda C_\theta \left(\frac{1}{2^n}\right)^{\theta+1} \right\} \leq C \left[ \lambda^{\frac{j}{p}} C_\theta^{\frac{j}{p}} \left(\frac{1}{2^n}\right)^{\frac{j}{p}(\theta+1)} \right]^{-2\tilde{N}} \|f_{m,n,k}^j\|_{q,N} \quad (4.3)$$

where  $C$  depends only on  $d, q$  and  $N$ . It thus follows that

$$I_j(m) \leq C \sum_{n=1}^{\infty} \sum_{k=1}^{2^n} \left[ \lambda^{\frac{j}{p}} C_\theta^{\frac{j}{p}} \left(\frac{1}{2^n}\right)^{\frac{j}{p}(\theta+1)} \right]^{-2\tilde{N}} \|f_{m,n,k}^j\|_{q,N}. \quad (4.4)$$

Therefore, we need to estimate the Sobolev norm  $\|f_{m,n,k}^j\|_{q,N}$  in order to prove our main theorem 2.3, and we will see that there is a good reason (see the constraint (4.18) below) that we need to raise the power of  $|w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j}|^{\frac{p}{j}}$  to  $2\tilde{N}$  for large enough  $\tilde{N}$ .

To this end, we first need to evaluate higher order Malliavin derivatives of  $f_{m,n,k}^j$ . For simplicity, let  $X_j(w) = w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j}$  and  $Y_j(w) = w_{t_n^{k-1}, t_n^k}^{(m),j}$  ( $j = 1, 2$ ). Suppose  $\tilde{N} \in \mathbb{N}$



is chosen, and consider  $f = |X|^{2\tilde{N}}$  where  $X = X_j$  or  $Y_j$  ( $j = 1, 2$ ). In all these cases,  $f$  is a polynomial of Brownian motion path, and thus is smooth in the Malliavin sense. In particular,  $f \in \mathbb{D}_N^q$  for any  $q \geq 1$  and  $N \in \mathbb{N}$ . We want to find an upper bound for the Sobolev norm  $\|f\|_{q,N}$ .

If  $M \leq \tilde{N}$ , we have

$$D^M f = \sum_{\mu=1}^M \sum_{\substack{\alpha_1+\dots+\alpha_\mu=M \\ 4 \geq \alpha_i \geq 1}} C_{\alpha_1 \dots \alpha_\mu} |X|^{2(\tilde{N}-\mu)} D^{\alpha_1} |X|^2 \otimes \dots \otimes D^{\alpha_\mu} |X|^2 \quad (4.5)$$

where  $C_{\alpha_1 \dots \alpha_\mu}$  are constants depending only on  $\alpha$ 's,  $j = 1$  or  $2$ ,  $M$  and  $\tilde{N}$ . Therefore, by using Hölder's inequality, we have

$$\begin{aligned} \|f\|_{q,N} &\leq C \sum_{M=0}^N \sum_{\mu=1}^M \sum_{\substack{\alpha_1+\dots+\alpha_\mu=M \\ 4 \geq \alpha_i \geq 0}} \| |X|^{2(\tilde{N}-\mu)} D^{\alpha_1} |X|^2 \otimes \dots \otimes D^{\alpha_\mu} |X|^2 \|_q \\ &\leq C \sum_{M=0}^N \sum_{\mu=1}^M \sum_{\substack{a_1+\dots+a_\mu=M \\ 4 \geq a_i \geq 0}} \| |X|^{2(\tilde{N}-\mu)} \|_{4q(\tilde{N}-\mu)} \prod_{i=1}^{\mu} \| D^{\alpha_i} |X|^2 \|_{2\mu q} \end{aligned} \quad (4.6)$$

for some constant  $C$  depending only on  $N, \tilde{N}$ , where the restriction that  $4 \geq a_i \geq 0$  comes from the fact that  $|X|^2$  is a polynomial of Brownian motion of order at most four as  $X = X_j$  or  $Y_j$ , so that  $D^a |X|^2 = 0$  for  $a \geq 5$ . The inequality (4.6), though completely elementary, allows us to develop the necessary estimates for the Sobolev norms we are interested.

**Lemma 4.1** *Suppose  $X$  is a smooth Malliavin functional, and  $D^a X = 0$  for  $a \geq 3$ , then*

$$|D|X|^2|_{\mathbf{H}} \leq 2|X||DX|_{\mathbf{H}}, \quad |D^2|X|^2|_{\mathbf{H}^{\otimes 2}} \leq 2|DX|_{\mathbf{H}}^2 + 2|X||D^2X|_{\mathbf{H}^{\otimes 2}} \quad (4.7)$$

and

$$|D^3|X|^2|_{\mathbf{H}^{\otimes 3}} \leq 6|D^2X|_{\mathbf{H}^{\otimes 2}}|DX|_{\mathbf{H}}, \quad |D^4|X|^2|_{\mathbf{H}^{\otimes 4}} \leq 6|D^2X|_{\mathbf{H}^{\otimes 2}}^2. \quad (4.8)$$

**Proof.** These estimates follows from the chain rule directly. ■

**Lemma 4.2** *Let  $m, n \in N$ , and  $k = 1, \dots, 2^n$ . Let  $Y_j = w_{t_n^{k-1}, t_n^k}^{(m),j}$  and  $X_j = w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j}$  ( $j = 1, 2$ ). Then, there is a constant  $C$  depending only on  $d$ , such that for any  $q \geq 1$*

$$\|D^a |Y_1|^2\|_q \leq \begin{cases} C\sqrt{q} \frac{2^m}{2^{2n}} & \text{if } n > m, \\ C\sqrt{q} \frac{1}{2^n} & \text{if } n \leq m, \end{cases} \quad (4.9)$$

$$\|D^a |X_1|^2\|_q \leq \begin{cases} C\sqrt{q} \frac{2^m}{2^{2n}} & \text{if } n > m, \\ 0 & \text{if } n \leq m \end{cases} \quad (4.10)$$

and

$$\|D^b |X_2|^2\|_q \leq \begin{cases} C\sqrt{q} \frac{4^{-b} 2^{2m}}{2^{4n}} & \text{for } n \geq m, \\ C\sqrt{q} \frac{1}{2^{m+n}} & \text{for } n < m \end{cases} \quad (4.11)$$

where  $a = 1, 2$  and  $b = 1, 2, 3, 4$ .

**Proof.** If  $n \geq m$ , then  $Y_1 = \frac{2^m}{2^n} \xi_m^{k(n,m)}$ ,  $DY_1 = \frac{2^m}{2^n} \mathbf{1}_{J_m^{k(n,m)}}$  and  $D^2Y_1 = 0$ , so that

$$|D|Y_1|^2|_{\mathbf{H}} \leq 2|Y_1||DY_1|_{\mathbf{H}} = 2 \frac{2^m}{2^n} \frac{2^m}{2^n} \sqrt{\frac{1}{2^m}} |\xi_m^{k(n,m)}| \quad (4.12)$$

which yields

$$\|D|Y_1|^2\|_q \leq C \sqrt{q} \frac{2^m}{2^{2n}}$$

where  $C$  depends only on  $d$ . Similarly

$$|D^2|Y_1|^2|_{\mathbf{H}^{\otimes 2}} \leq 2|DY_1|_{\mathbf{H}}^2 = 2 \frac{2^{2m}}{2^{2n}} \mathbf{1}_{J_m^{k(n,m)}}$$

so that

$$\|D^2|Y_1|^2\|_q \leq 2 \frac{2^m}{2^{2n}}.$$

If  $n < m$ , then  $Y_1 = \xi_n^k$  so that  $DY_1 = \mathbf{1}_{J_n^k}$ , hence

$$\|D|Y_1|^2\|_q \leq C \sqrt{q} \frac{1}{2^n}, \quad \|D^2|Y_1|^2\|_q \leq C \frac{1}{2^n}$$

where  $C$  depends only on  $d$ . This proves (4.9). (4.10) follows (4.9) and the fact that  $X_1 = 0$  if  $n \leq m$ .

Together with Lemma 4.1 and the  $L^q$ -estimate (3.14) for  $X_2$ , we can conclude that there is a constant  $C$  depending only on  $d$  such that

$$\|D^\alpha |X_2|^2\|_q \leq C \sqrt{q}^{4-\alpha} \frac{1}{2^{m+n}} \quad \forall n < m \quad (4.13)$$

for  $\alpha = 1, 2, 3, 4$  and  $q \geq 1$ .

Now consider the case that  $n > m$ . In this case

$$Y_2 = w_{t_n^{k-1}, t_n^k}^{(m),2} = \frac{1}{2} \frac{2^{2m}}{2^{2n}} \xi_m^{k(n,m)} \otimes \xi_m^{k(n,m)}$$

so that

$$DY_2 = \frac{1}{2} \frac{2^{2m}}{2^{2n}} \{\mathbf{1}_{J_m^{k(n,m)}}, \xi_m^{k(n,m)}\}, \quad D^2Y_2 = \frac{1}{2} \frac{2^{2m}}{2^{2n}} \{\mathbf{1}_{J_m^{k(n,m)}}, \mathbf{1}_{J_m^{k(n,m)}}\}.$$

where

$$\{\mathbf{1}_{J_m^{k(n,m)}}, \mathbf{1}_{J_m^{k(n,m)}}\}(t_1, t_2) = \{\mathbf{1}_{J_m^{k(n,m)}}(t_1), \mathbf{1}_{J_m^{k(n,m)}}(t_2)\}.$$

It follows that

$$\|D^b |Y_2|^2\|_q \leq C q^{4-b} \frac{2^{2m}}{2^{4n}} \quad \text{for } n > m, b = 1, 2, 3, 4.$$

and therefore (4.11). ■

In what follows we assume that  $\tilde{N} \in \mathbb{N}$  and  $N \leq \tilde{N}$ . Let

$$f_{m,n,k}^j(w) = \left| w_{\frac{k-1}{2^n}, \frac{k}{2^n}}^{(m+1),j} - w_{\frac{k-1}{2^n}, \frac{k}{2^n}}^{(m),j} \right|^{2\tilde{N}}, \quad g_{m,n,k}^j(w) = \left| w_{\frac{k-1}{2^n}, \frac{k}{2^n}}^{(m),j} \right|^{2\tilde{N}}. \quad (4.14)$$

**Lemma 4.3** *There is a constant  $C$  depending only on  $N, \tilde{N}$  and  $d$  such that*

$$\|f_{m,n,k}^1\|_{q,N} \leq Cq^{\tilde{N}} \left(\frac{2^m}{2^{2n}}\right)^{\tilde{N}} \quad \text{for } n > m \quad (4.15)$$

and

$$\|g_{m,n,k}^1\|_{q,N} \leq \begin{cases} Cq^{\tilde{N}} \left(\frac{2^m}{2^{2n}}\right)^{\tilde{N}} & \text{for } n > m \\ Cq^{\tilde{N}} \left(\frac{1}{2^n}\right)^{\tilde{N}} & \text{for } n \leq m \end{cases} \quad (4.16)$$

for all  $q \geq 1$ .

**Proof.** If  $X_1 = w_{\frac{k-1}{2^n}, \frac{k}{2^n}}^{(m+1),1} - w_{\frac{k-1}{2^n}, \frac{k}{2^n}}^{(m),1}$  (for  $n > m$ , otherwise  $X_1 = 0$ ). By (4.6)

$$\begin{aligned} \|f_{m,n,k}^1\|_{q,N} &\leq C \sum_{M=0}^N \sum_{\mu=1}^M \sum_{\substack{a_1+\dots+a_\mu=M \\ 2 \geq a_i \geq 0}} \|X_1\|_{4q(\tilde{N}-\mu)}^{2(\tilde{N}-\mu)} \prod_{i=1}^{\mu} \|D^{\alpha_i} |X_1|^2\|_{2\mu q} \\ &\leq C \sum_{M=0}^N \sum_{\mu=1}^M \sum_{\substack{a_1+\dots+a_\mu=M \\ 4 \geq a_i \geq 0}} \left(\sqrt{q} \frac{2^m}{2^{2n}}\right)^{\mu} \left(q \frac{2^m}{2^{2n}}\right)^{(\tilde{N}-\mu)} \\ &\leq Cq^{\tilde{N}} \left(\frac{2^m}{2^{2n}}\right)^{\tilde{N}} \quad \text{for } n > m. \end{aligned}$$

The same estimate remains true in the case that  $Y_1 = w_{\frac{k-1}{2^n}, \frac{k}{2^n}}^{(m),1}$  and  $n > m$ . On the other hand, if  $n \leq m$ , then

$$\begin{aligned} \|g_{m,n,k}^1\|_{q,N} &\leq C \sum_{M=0}^N \sum_{\mu=1}^M \sum_{\substack{a_1+\dots+a_\mu=M \\ 2 \geq a_i \geq 0}} \|Y_1\|_{4q(\tilde{N}-\mu)}^{2(\tilde{N}-\mu)} \prod_{i=1}^{\mu} \|D^{\alpha_i} |Y_1|^2\|_{2\mu q} \\ &\leq Cq^{\tilde{N}} \sum_{M=0}^N \sum_{\mu=1}^M \sum_{\substack{a_1+\dots+a_\mu=M \\ 2 \geq a_i \geq 0}} \left(\frac{1}{2^n}\right)^{\mu} \left(\frac{1}{2^n}\right)^{(\tilde{N}-\mu)} \\ &= Cq^{\tilde{N}} \left(\frac{1}{2^n}\right)^{\tilde{N}} \quad \text{for } n \leq m. \end{aligned}$$

■

**Lemma 4.4** *There is a constant  $C$  depending only on  $N, \tilde{N}, d$  such that*

$$\|f_{m,n,k}^2\|_{q,N} \leq \begin{cases} C\sqrt{q}^{4\tilde{N}} \left(\frac{1}{2^n}\right)^{\tilde{N}} \left(\frac{1}{2^m}\right)^{\tilde{N}} & \text{for } n \leq m, \\ Cq^{4\tilde{N}} \left(\frac{2^m}{2^{2n}}\right)^{2\tilde{N}} & \text{for } n > m \end{cases} \quad (4.17)$$

for all  $q \geq 1$ .

**Proof.** Let  $X_2 = w_{\frac{k-1}{2^n}, \frac{k}{2^n}}^{(m+1),j} - w_{\frac{k-1}{2^n}, \frac{k}{2^n}}^{(m),j}$ . By (4.6) and the  $L^q$ -bounds of  $X_2$  applying to  $4q(\tilde{N} - \mu)$  and (4.11)

$$\|X_2\|_{4q(\tilde{N}-\mu)} \leq \begin{cases} Cq^{\frac{2^m}{2^{2n}}} & \text{if } n > m \\ Cq\sqrt{\frac{1}{2^{m+n}}} & \text{if } n \leq m, \end{cases}$$

and

$$\|D^b|X_2|^2\|_{2\mu q} \leq \begin{cases} Cq^2\frac{2^{2m}}{2^{4n}} & \text{for } n > m, \\ Cq^2\frac{1}{2^{m+n}} & \text{for } n \leq m, \end{cases}$$

we obtain, for  $n \leq m$ ,

$$\begin{aligned} \|f_{m,n,k}^2\|_{q,N} &\leq C \sum_{M=0}^N \sum_{\mu=1}^M \sum_{\substack{a_1+\dots+a_\mu=M \\ 4 \geq a_i \geq 0}} \left( q\sqrt{\frac{1}{2^{m+n}}} \right)^{2(\tilde{N}-\mu)} \left( q^2\frac{1}{2^{m+n}} \right)^\mu \\ &\leq C\sqrt{q}^{4\tilde{N}} \left( \frac{1}{2^{m+n}} \right)^{\tilde{N}}. \end{aligned}$$

Similarly, if  $n > m$  then

$$\begin{aligned} \|f_{m,n,k}^2\|_{q,N} &\leq C \sum_{M=0}^N \sum_{\mu=1}^M \sum_{\substack{a_1+\dots+a_\mu=M \\ 4 \geq a_i \geq 0}} \left( q\frac{2^m}{2^{2n}} \right)^{2(\tilde{N}-\mu)} \left( q^2\frac{2^{2m}}{2^{4n}} \right)^\mu \\ &\leq C\sqrt{q}^{4\tilde{N}} \left( \frac{2^m}{2^{2n}} \right)^{2\tilde{N}}. \end{aligned}$$

■

**Proposition 4.5** Choose  $\tilde{N} \in \mathbb{N}$  and  $\theta, \beta > 0$  such that

$$p - 2 - \frac{p}{\tilde{N}} > 0, \quad \beta + \theta < \frac{p - 2 - \frac{p}{\tilde{N}}}{2} \quad (4.18)$$

Then for any  $N \leq \tilde{N}$  there is a constant  $C$  depending only on  $N, \tilde{N}, \theta, \beta, d$  and  $q \geq 1$  such that

$$Cap_{q,N} \left\{ \sum_{k=1}^{2^n} \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{p}{j}} > C_\theta \left( \frac{1}{2^m} \right)^\beta \left( \frac{1}{2^n} \right)^\theta \right\} \leq C \left( \frac{1}{2^{\max\{m,n\}}} \right)^{\varepsilon_j} \quad (4.19)$$

for  $n, m \in \mathbb{N}$ ,  $k = 1, \dots, 2^n$ , where

$$\varepsilon_j = \left[ \frac{p-2}{2} - (\theta + \beta) \right] \frac{2j\tilde{N}}{p} - 1, \quad j = 1, 2. \quad (4.20)$$

**Proof.** For each fixed  $k = 1, \dots, 2^n$ , consider  $f_{m,n,k}^j(w) = \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{2\tilde{N}}$  which is continuous on  $\mathbf{W}$ . Thus, according to the capacity maximal inequality

$$\begin{aligned} & Cap_{q,N} \left\{ \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{2}{j}} > C_\theta \left( \frac{1}{2^m} \right)^\beta \left( \frac{1}{2^n} \right)^{\theta+1} \right\} \\ &= Cap_{q,N} \left\{ f_{m,n,k}^j > \left[ C_\theta^{\frac{1}{p}} \left( \frac{1}{2^m} \right)^{\beta \frac{1}{p}} \left( \frac{1}{2^n} \right)^{\frac{1}{p}(\theta+1)} \right]^{2\tilde{N}} \right\} \\ &\leq C \left[ C_\theta^{\frac{1}{p}} \left( \frac{1}{2^m} \right)^{\beta \frac{1}{p}} \left( \frac{1}{2^n} \right)^{\frac{1}{p}(\theta+1)} \right]^{-2\tilde{N}} \|f_{m,n,k}^j\|_{q,N}. \end{aligned}$$

On the other hand

$$\begin{aligned} & Cap_{q,N} \left\{ \sum_{k=1}^{2^n} \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{2}{j}} > C_\theta \left( \frac{1}{2^m} \right)^\beta \left( \frac{1}{2^n} \right)^\theta \right\} \\ &\leq \sum_{k=1}^{2^n} Cap_{q,N} \left\{ \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|^{\frac{2}{j}} > C_\theta \left( \frac{1}{2^m} \right)^\beta \left( \frac{1}{2^n} \right)^{\theta+1} \right\} \end{aligned}$$

It follows from (4.17) that

$$\begin{aligned} & Cap_{q,N} \left\{ \sum_{k=1}^{2^n} \left| w_{t_n^{k-1}, t_n^k}^{(m+1),2} - w_{t_n^{k-1}, t_n^k}^{(m),2} \right|^{\frac{2}{2}} > C_\theta \left( \frac{1}{2^m} \right)^\beta \left( \frac{1}{2^n} \right)^\theta \right\} \\ &\leq \begin{cases} C 2^n \left[ \left( \frac{1}{2^m} \right)^{\frac{2}{p}\beta} C_\theta^{\frac{2}{p}} \left( \frac{1}{2^n} \right)^{\frac{2}{p}(\theta+1)} \right]^{-2\tilde{N}} \left( \frac{1}{2^n} \right)^{\tilde{N}} \left( \frac{1}{2^m} \right)^{\tilde{N}} & \text{for } n \leq m, \\ C 2^n \left[ \left( \frac{1}{2^m} \right)^{\frac{2}{p}\beta} C_\theta^{\frac{2}{p}} \left( \frac{1}{2^n} \right)^{\frac{2}{p}(\theta+1)} \right]^{-2\tilde{N}} \left( \frac{2^m}{2^{2n}} \right)^{2\tilde{N}} & \text{for } n > m \end{cases} \\ &\leq C \left( \frac{1}{2^{m \vee n}} \right)^{\frac{p-2}{p} 2\tilde{N} - \frac{4}{p} \tilde{N}(\beta+\theta)-1} \quad \text{for all } n \text{ and } m. \end{aligned}$$

Similarly, for  $j = 1$  and  $n > m$  we have

$$\begin{aligned} & Cap_{q,N} \left\{ \sum_{k=1}^{2^n} \left| w_{t_n^{k-1}, t_n^k}^{(m+1),1} - w_{t_n^{k-1}, t_n^k}^{(m),1} \right|^p > C_\theta \left( \frac{1}{2^m} \right)^\beta \left( \frac{1}{2^n} \right)^\theta \right\} \\ &\leq C 2^n \left[ \left( \frac{1}{2^m} \right)^{\frac{1}{p}\beta} C_\theta^{\frac{1}{p}} \left( \frac{1}{2^n} \right)^{\frac{1}{p}(\theta+1)} \right]^{-2\tilde{N}} \|f_1\|_{q,N} \\ &\leq C 2^n \left[ \left( \frac{1}{2^m} \right)^{\frac{1}{p}\beta} C_\theta^{\frac{1}{p}} \left( \frac{1}{2^n} \right)^{\frac{1}{p}(\theta+1)} \right]^{-2\tilde{N}} \left( \frac{2^m}{2^{2n}} \right)^{\tilde{N}} \\ &\leq C \left( \frac{1}{2^n} \right)^{\frac{p-2}{p} \tilde{N} - \frac{2}{p} \tilde{N}(\theta+\beta)-1} \end{aligned}$$

which completes the proof. ■

**Proposition 4.6** *Let  $p \in (2, 3)$ ,  $q \geq 1$ ,  $N \in \mathbb{N}$ , and  $\beta \in (0, \frac{p-2}{2})$ . Then, for any  $\varepsilon > 0$  there is a constant  $C$  depending only on  $p, d, q, N$  and  $\beta$  such that*

$$Cap_{q,N} \left\{ \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) > \left( \frac{1}{2^m} \right)^{\frac{j}{p}\beta} \right\} \leq C \left( \frac{1}{2^m} \right)^\varepsilon \quad (4.21)$$

for all  $m \in \mathbb{N}$  and  $j = 1, 2$ .

**Proof.** Choose  $\theta > 0$  and  $\tilde{N} \in \mathbb{N}$  such that  $\tilde{N} > N$ ,

$$p - 2 - \frac{p}{\tilde{N}} > 0, \quad \beta + \theta < \frac{p-2}{2} - \frac{p}{2\tilde{N}}$$

and

$$\left[ \frac{p-2}{2} - (\theta + \beta) \right] \frac{2\tilde{N}}{p} - 1 \geq 2\varepsilon.$$

Then according to Proposition 4.5, there is a constant  $C$  depending only on  $N, \tilde{N}, \theta, \beta, d$  and  $q$  such that

$$\begin{aligned} & Cap_{q,\alpha} \left\{ \rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})_{\frac{p}{j}} > \left( \frac{1}{2^m} \right)^\beta \right\} \\ & \leq \sum_{n=1}^{\infty} Cap_{q,N} \left\{ \sum_{k=1}^{2^n} \left| w_{t_n^{k-1}, t_n^k}^{(m+1),j} - w_{t_n^{k-1}, t_n^k}^{(m),j} \right|_{\frac{p}{j}} > C_\theta \left( \frac{1}{2^m} \right)^\beta \left( \frac{1}{2^n} \right)^\theta \right\} \\ & \leq C \sum_{n=1}^m \left( \frac{1}{2^m} \right)^{2\varepsilon} + C \sum_{n \geq m} \left( \frac{1}{2^n} \right)^{2\varepsilon} \\ & \leq C \left( \frac{1}{2^m} \right)^\varepsilon. \end{aligned}$$

■

This proposition shows the capacity of  $\{\rho_j(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) > 2^{-j\beta m/p}\}$  for small  $\beta > 0$  decays sub-exponentially in  $2^{-m}$  (in contrast with the decay rate in (3.29) which is indeed not a sharp estimate). This is the right order for  $j = 2$ . In the case  $j = 1$  and for capacity  $Cap_{2,1}$ , this result was established by M. Fukushima [16].

**Lemma 4.7** *Let  $p \in (2, 3)$ ,  $q \geq 1$ ,  $N \in \mathbb{N}$ ,  $\delta > 0$  and  $\tilde{N} \in \mathbb{N}$  such that*

$$N \leq \tilde{N}, \quad \tilde{N} \left( 1 - \frac{2}{p} \right) - 1 > 0,$$

there is a constant  $C$  depending only on  $N, p, d, \delta, q \geq 1$  such that

$$Cap_{q,N} \left\{ \rho_1(\mathbf{w}^{(m)}) > (2^m)^{\frac{\delta}{p}} \right\} \leq C \left( \frac{1}{2^m} \right)^{\frac{2\delta}{p}\tilde{N}} \quad \forall m \in \mathbb{N}. \quad (4.22)$$

**Proof.** Choose  $\theta > 0$  such that

$$\tilde{N} \left( 1 - 2^{\frac{\theta+1}{p}} \right) - 1 > 0.$$

Then

$$\begin{aligned} \text{Cap}_{q,N} \left\{ \rho_1(\mathbf{w}^{(m)})^p > (2^m)^\delta \right\} &\leq \sum_{n=1}^{\infty} \text{Cap}_{q,N} \left\{ \sum_{k=1}^{2^n} \left| w_{\frac{k-1}{2^n}, \frac{k}{2^n}}^{(m),1} \right|^p > C_\theta (2^m)^\delta \left( \frac{1}{2^n} \right)^\theta \right\} \\ &\leq \sum_{n=1}^{\infty} \sum_{k=1}^{2^n} \text{Cap}_{q,N} \left\{ \left| w_{\frac{k-1}{2^n}, \frac{k}{2^n}}^{(m),1} \right|^p > C_\theta (2^m)^\delta \left( \frac{1}{2^n} \right)^{\theta+1} \right\} \\ &\leq \sum_{n=1}^{\infty} \sum_{k=1}^{2^n} \text{Cap}_{q,N} \left\{ g_{m,n,k}^1 > \left[ C_\theta^{\frac{1}{p}} (2^m)^{\frac{\delta}{p}} \left( \frac{1}{2^n} \right)^{\frac{\theta+1}{p}} \right]^{2\tilde{N}} \right\} \end{aligned}$$

where

$$g_{m,n,k}^1 = \left| w_{\frac{k-1}{2^n}, \frac{k}{2^n}}^{(m),1} \right|^{2\tilde{N}}.$$

Thus, by using the capacity maximal inequality and (4.16):

$$\|g_{m,n,k}^1\|_{q,N} \leq \begin{cases} C q^{\tilde{N}} \left( \frac{2^m}{2^{2n}} \right)^{\tilde{N}} & \text{for } n > m, \\ C q^{\tilde{N}} \left( \frac{1}{2^n} \right)^{\tilde{N}} & \text{for } n \leq m, \end{cases}$$

we obtain

$$\begin{aligned} \text{Cap}_{q,N} \left\{ \rho_1(\mathbf{w}^{(m)})^p > (2^m)^\delta \right\} &\leq C \sum_{n=1}^m 2^n \left[ C_\theta^{\frac{1}{p}} (2^m)^{\frac{\delta}{p}} \left( \frac{1}{2^n} \right)^{\frac{\theta+1}{p}} \right]^{-2\tilde{N}} \left( \frac{1}{2^n} \right)^{\tilde{N}} \\ &\quad + C \sum_{n>m} 2^n \left[ C_\theta^{\frac{1}{p}} (2^m)^{\frac{\delta}{p}} \left( \frac{1}{2^n} \right)^{\frac{\theta+1}{p}} \right]^{-2\tilde{N}} \left( \frac{2^m}{2^{2n}} \right)^{\tilde{N}} \\ &\leq C \left( \frac{1}{2^m} \right)^{\tilde{N} \frac{2\delta}{p}} \sum_{n=1}^m \left( \frac{1}{2^n} \right)^{\tilde{N} (1 - 2^{\frac{\theta+1}{p}}) - 1} \\ &\quad + C \left( \frac{1}{2^m} \right)^{\tilde{N} \frac{2\delta}{p}} \sum_{n>m} \left( \frac{2^m}{2^n} \right)^{\tilde{N}} \left( \frac{1}{2^n} \right)^{\tilde{N} (1 - 2^{\frac{\theta+1}{p}}) - 1} \\ &\leq C \left( \frac{1}{2^m} \right)^{\tilde{N} \frac{2\delta}{p}}. \end{aligned}$$

■

**Proposition 4.8** *Let  $p \in (2, 3)$ ,  $q \geq 1$ ,  $N \in \mathbb{N}$ , and  $\beta \in (0, \frac{p-2}{2})$ . Then, for any  $\varepsilon > 0$  there is a constant  $C$  depending only on  $p, d, q, N$  and  $\beta$  such that*

$$\text{Cap}_{q,N} \left\{ \rho_1(\mathbf{w}^{(m)}) \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) > \left( \frac{1}{2^m} \right)^{\frac{\beta}{p}} \right\} \leq C \left( \frac{1}{2^m} \right)^\varepsilon \quad (4.23)$$

for all  $m \in \mathbb{N}$ .

**Proof.** Choose  $\tilde{N} \in \mathbb{N}$ ,  $\theta > 0$ ,  $\delta > 0$  such that

$$\tilde{N} \left( 1 - 2 \frac{\theta + 1}{p} \right) - 1 > \varepsilon,$$

$$p - 2 - \frac{p}{\tilde{N}} > 0, \beta + \theta + \delta < \frac{p-2}{2} - \frac{p}{2\tilde{N}}$$

and

$$\left[ \frac{p-2}{2} - (\theta + \beta + \delta) \right] \frac{2\tilde{N}}{p} - 1 \geq 2\varepsilon.$$

Then, since

$$\begin{aligned} & \left\{ \rho_1(\mathbf{w}^{(m)})^p \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p > \left( \frac{1}{2^m} \right)^\beta \right\} \\ \subset & \left\{ \rho_1(\mathbf{w}^{(m)})^p > (2^m)^\delta \right\} \cup \left\{ \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p > \left( \frac{1}{2^m} \right)^{\beta+\delta} \right\} \end{aligned}$$

so that

$$\begin{aligned} & Cap_{q,N} \left\{ \rho_1(\mathbf{w}^{(m)})^p \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p > \left( \frac{1}{2^m} \right)^\beta \right\} \\ \leq & Cap_{q,N} \left\{ \rho_1(\mathbf{w}^{(m)})^p > (2^m)^\delta \right\} + Cap_{q,N} \left\{ \rho_1(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)})^p > \left( \frac{1}{2^m} \right)^{\beta+\delta} \right\} \\ \leq & C \left( \frac{1}{2^m} \right)^\varepsilon. \end{aligned}$$

■

Putting ((2.11), (4.21) and (4.23) together we may conclude the following

**Theorem 4.9** *Let  $p \in (2, 3)$ . Then for any  $\varepsilon > 0$ ,  $q \geq 1$  and  $N \in \mathbb{N}$ , there are  $\beta > 0$ , constants  $C_1 > 0$  and  $C_2 > 0$  depending only on  $p, q, N, d$  and  $\varepsilon$  such that*

$$Cap_{q,N} \left\{ d_p(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) > C_1 \left( \frac{1}{2^m} \right)^\beta \right\} \leq C_2 \left( \frac{1}{2^m} \right)^\varepsilon \quad \forall m \in \mathbb{N}. \quad (4.24)$$

We are in a position to prove the main theorem 2.3. By the capacity version of the Borel-Cantelli lemma, (4.24) implies that

$$A = \left\{ w \in \mathbf{W} : \sum_{m=1}^{\infty} d_p(\mathbf{w}^{(m+1)}, \mathbf{w}^{(m)}) = \infty \right\}$$

is slim, that is,  $Cap_{q,N} \{A\} = 0$  for any  $q \geq 1$  and  $N \in \mathbb{N}$ , so that

$$\{w \in \mathbf{W} : (\mathbf{w}^{(m)}) \text{ is not Cauchy in } G_p\Omega(\mathbb{R}^d)\}$$

is slim, and therefore  $\mathbf{w}^{(m)} \rightarrow \mathbf{w}$  in  $G_p\Omega(\mathbb{R}^d)$  quasi-surely.



# References

- [1] R. F. Bass, B. M. Hambly, and T. J. Lyons, *Extending the Wong-Zakai theorem to reversible Markov processes*, J. Eur. Math. Soc. (JEMS) **4** (2002), no. 3, 237–269.
- [2] H. Bessaih, M. Gubinelli and F. Russo, *The evolution of a random vortex filament*, Ann. Probab. **33** (2005), 1825–1855.
- [3] N. Bouleau and F. Hirsch, *Dirichlet forms and analysis on Wiener space*, de Gruyter Studies in Mathematics, vol. 14, Walter de Gruyter & Co., Berlin, 1991.
- [4] M. Caruana and P. Friz, *Partial differential equations driven by rough paths*, J. Differential Equations **247** (2009) No.1, 140–173.
- [5] M. Caruana, P. Friz and H. Oberhauser, *A (rough) pathwise approach to a class of non-linear stochastic partial differential equations*, Ann. Inst. H. Poincaré Anal. Non Linéaire **28** (2011) no 1, 27–46.
- [6] L. Coutin and Z. Qian, *Stochastic analysis, rough path analysis and fractional Brownian motions*, Probab. Theory Related Fields **122** (2002), no. 1, 108–140.
- [7] M. Davie, *Differential equations driven by rough paths: an approach via discrete approximation*, Appl. Math. Res. Express. AMRX **2** (2007), Art. ID abm009, 40.
- [8] A. M. Davie, *Uniqueness of solutions of stochastic differential equations*, Int. Math. Res. Not. IMRN **24** (2007), Art. ID rnm124, 26.
- [9] A. Deya and S. Tindel, *Rough Volterra equations. I. The algebraic integration setting*, Stoch. Dyn. **9** (2009) no 3, 437–477.
- [10] S. Fang, *Sur la continuité approximative forte des fonctionnelles d’Itô*, Stochastics Stochastics Rep. **36** (1991), no. 3–4, 193–204.
- [11] S. Fang, *Analyse quasi-sure et l’estimation du noyau de la chaleur pour temps petit*, Trans. Amer. Math. Soc. **339** (1993), no. 1, 221–241.
- [12] P. K. Friz and N. B. Victoir, *Multidimensional stochastic processes as rough paths*, Cambridge Studies in Advanced Mathematics, vol. 120, Cambridge University Press, Cambridge, 2010, Theory and applications.
- [13] M. Fukushima, *A note on irreducibility and ergodicity of symmetric Markov processes*, Stochastic processes in quantum theory and statistical physics (Marseille, 1981), Lecture Notes in Phys., vol. 173, Springer, Berlin, 1982, pp. 200–207.
- [14] M. Fukushima, *Dirichlet forms and Markov processes*, North-Holland Mathematical Library, vol. 23, North-Holland Publishing Co., Amsterdam, 1980.
- [15] M. Fukushima, *Capacitary maximal inequalities and an ergodic theorem*, Probability theory and mathematical statistics (Tbilisi, 1982), Lecture Notes in Math., vol. 1021, Springer, Berlin, 1983, pp. 130–136.

- [16] M. Fukushima, *Basic properties of Brownian motion and a capacity on the Wiener space*, J. Math. Soc. Japan **36** (1984), no. 1, 161–176.
- [17] M. Fukushima, *A note on capacities in infinite dimensions*, Probability theory and mathematical statistics (Kyoto, 1986), Lecture Notes in Math., vol. 1299, Springer, Berlin, 1988, pp. 80–85.
- [18] M. Fukushima,  *$(r, p)$ -capacities and Hunt processes in infinite dimensions*, Probability theory and mathematical statistics (Kiev, 1991), World Sci. Publ., River Edge, NJ, 1992, pp. 96–103.
- [19] M. Gubinelli, and A. Lejay and S. Tindel, *Young integrals and SPDEs*, Potential Anal. **25** (2006) no 4, 307–326.
- [20] M. Hairer and H. Weber, *Rough Burgers-like equations with multiplicative noise*, ArXiv e-prints (2010), available at [1011.0966](https://arxiv.org/abs/1011.0966).
- [21] B. M. Hambly and T. J. Lyons, *Stochastic area for Brownian motion on the Sierpinski gasket*, Ann. Probab. **26** (1998), no. 1, 132–148.
- [22] F. Hirsch and S. Song, *Propriétés de Markov des processus à plusieurs paramètres et capacités*, C. R. Acad. Sci. Paris Sér. I Math. **319** (1994), no. 5, 483–488.
- [23] Z. Huang and J. Ren, *Quasi sure stochastic flows*, Stochastics Stochastics Rep. **33** (1990), no. 3–4, 149–157.
- [24] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Amsterdam, 1981.
- [25] K. Itô, *An elementary approach to Malliavin fields*, Asymptotic problems in probability theory: Wiener functionals and asymptotics (Sanda/Kyoto, 1990), Pitman Res. Notes Math. Ser., vol. 284, Longman Sci. Tech., Harlow, 1993, pp. 35–89.
- [26] K. Itô and H. P. McKean, Jr., *Diffusion processes and their sample paths*, Springer-Verlag, Berlin, 1974, Second printing, corrected, Die Grundlehren der mathematischen Wissenschaften, Band 125.
- [27] H. Kaneko, *On  $(r, p)$ -capacities for Markov processes*, Osaka J. Math. **23** (1986), no. 2, 325–336.
- [28] T. Kazumi, *Refinement in terms of capacities of certain limit theorems on an abstract Wiener space*, J. Math. Kyoto Univ. **32** (1992), no. 1, 1–29.
- [29] S. Kusuoka, *Analysis on Wiener spaces. I. Nonlinear maps*, J. Funct. Anal. **98** (1991), no. 1, 122–168.
- [30] S. Kusuoka, *Analysis on Wiener spaces. II. Differential forms*, J. Funct. Anal. **103** (1992), no. 2, 229–274.

- [31] M. Ledoux, Z. Qian, and T. Zhang, *Large deviations and support theorem for diffusion processes via rough paths*, Stochastic Process. Appl. **102** (2002), no. 2, 265–283.
- [32] A. Lejay, *An introduction to rough paths*, Séminaire de Probabilités XXXVII, Lecture Notes in Math., vol. 1832, Springer, Berlin, 2003, pp. 1–59.
- [33] P. Lévy, *Processus stochastiques et mouvement brownien*, Paris, 1948.
- [34] P. Lévy, *Le mouvement brownien*, Paris, 1954.
- [35] T. Lyons and Z. Qian, *System control and rough paths*, Oxford Mathematical Monographs, Oxford University Press, Oxford, 2002, Oxford Science Publications.
- [36] T. Lyons, *Differential equations driven by rough signals*, Rev. Mat. Iberoamericana **14** (1998), no. 2, 215–310.
- [37] T. Lyons, M. Caruana, and T. Lévy, *Differential equations driven by rough paths*, Lecture Notes in Mathematics, vol. 1908, Springer, Berlin, 2007, Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, With an introduction concerning the Summer School by Jean Picard.
- [38] Z. Ma and M. Röckner, *Introduction to the theory of (nonsymmetric) Dirichlet forms*, Universitext, Springer-Verlag, Berlin, 1992.
- [39] P. Malliavin, *Calcul des variations, intégrales stochastiques et complexe de de Rham sur l'espace de Wiener*, C. R. Acad. Sci. Paris Sér. I Math. **299** (1984), no. 8, 347–350.
- [40] P. Malliavin, *Stochastic analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 313, Springer-Verlag, Berlin, 1997.
- [41] P. Malliavin and D. Nualart, *Quasi-sure analysis and Stratonovich anticipative stochastic differential equations*, Probab. Theory Related Fields **96** (1993), no. 1, 45–55.
- [42] E. J. McShane, *Stochastic differential equations and models of random processes*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory (Berkeley, Calif.), Univ. California Press, 1972, pp. 263–294.
- [43] P. A. Meyer, *Note sur les processus d'ornstein-uhlenbeck (appendice: Un resultat de d. williams)*, Séminaire de Probabilités XVI, Lecture Notes in Math., vol. 920, Springer, Berlin, 1982, pp. 95–132.
- [44] J. Ren, *Analyse quasi-sûre des équations différentielles stochastiques*, Bull. Sci. Math. **114** (1990), no. 2, 187–213.
- [45] J. Ren, *Topologie  $p$ -fine sur l'espace de Wiener et théorème des fonctions implicites*, Bull. Sci. Math. **114** (1990), no. 2, 99–114.

- [46] I. Shigekawa, *Sobolev spaces of Banach-valued functions associated with a Markov process*, Probab. Theory Related Fields **99** (1994), no. 3, 425–441.
- [47] D. W. Stroock and S. R. S. Varadhan, *On the support of diffusion processes with applications to the strong maximum principle*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory (Berkeley, Calif.), Univ. California Press, 1972, pp. 333–359.
- [48] H. Sugita, *Positive generalized Wiener functions and potential theory over abstract Wiener spaces*, Osaka J. Math. **25** (1988), no. 3, 665–696.
- [49] M. Takeda,  *$(r, p)$ -capacity on the Wiener space and properties of Brownian motion*, Z. Wahrsch. Verw. Gebiete **68** (1984), no. 2, 149–162.
- [50] E. Wong and M. Zakai, *On the relation between ordinary and stochastic differential equations*, Internat. J. Engrg. Sci. **3** (1965), 213–229.